

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Application of: Ryan

Serial No.: 10/608,463

Group Art Unit: 1652

Filed: June 27, 2003

Examiner: T. Saidha

FOR: ISOLATED GENOMIC POLYNUCLEOTIDE FRAGMENTS FROM  
CHROMOSOME 12 THAT ENCODE HUMAN CARBOXYPEPTIDASE M AND THE  
HUMAN MOUSE DOUBLE MINUTE 2 HOMOLOG

Confirmation No.: 6428

**APPEAL BRIEF**

Mail Stop Appeal  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

This Appeal Brief is being submitted in support of the appeal from the rejections set forth in the Office Action mailed on January 2, 2009.

## TABLE OF CONTENTS

I.	REAL PARTY IN INTEREST .....	3
II.	RELATED APPEALS AND INTERFERENCES .....	3
III.	STATUS OF CLAIMS .....	3
IV.	STATUS OF AMENDMENTS .....	3
V.	SUMMARY OF THE CLAIMED SUBJECT MATTER .....	3
VI.	GROUND OF REJECTION TO BE REVIEWED ON APPEAL.....	12
VII.	ARGUMENT .....	12
	A. Claims 7, 10, 15-18, 20, 24, 25, 30 and 31 Comply With the Written Description Requirement .....	12
	B. Claims 7, 10, 15-18, 20, 24, 25, 30 and 31 are not Obvious over Muzny et al. in view of Vogelstein .....	16
	(1) One of ordinary skill in the art could not have predicted where the MDM2 gene was located .....	18
	(2) The conflicting teachings in the art with respect to the possible location of the MDM2 gene is of particular significance .....	20
	(3) At best, this would still be an “obvious to try” situation .....	22
	(4) There was a long felt need and it is likely that others had tried but failed before the priority date of the instant application.....	24
	CLAIMS APPENDIX .....	26
	EVIDENCE APPENDIX .....	41
	RELATED PROCEEDINGS APPENDIX .....	42

## **I. REAL PARTY IN INTEREST**

The real party in interest of the present application is Ryogen LLC. Ryogen LLC is the owner of the present application by way of an assignment from the inventor, James W. Ryan of all rights, title, and interest.

## **II. RELATED APPEALS AND INTERFERENCES**

There are no appeals or interferences related to the present application.

## **III. STATUS OF CLAIMS**

Claims 7, 10, 15-18, 20, 24, 25, 30 and 31 stand finally rejected by the Examiner as noted in the Office Action mailed on January 2, 2009 and in the Advisory Action mailed June 24, 2009. Claims 1-6, 8-9, 11, 13, 19, 21 and 26-29 have been canceled. Claims 12, 14, 22, 23 and 32-38 have been withdrawn. The rejection of claims 7, 10, 15-18, 20, 24, 25, 30 and 31 is appealed.

## **IV. STATUS OF AMENDMENTS**

Amendments to claims 7 and 24 subsequent to the Examiner's Final Rejection mailed January 2, 2009 were submitted but were not entered.

## **V. SUMMARY OF THE CLAIMED SUBJECT MATTER**

### **A. Independent Claim 1**

Claim Elements	Support in specification
An isolated nucleic acid molecule 20-51039 contiguous nucleotides in length consisting of a reverse or forward strand of a region of SEQ ID NO:4	Page 9, line 34; page 10, lines 22-26
wherein said region is selected from the group consisting of a 5'-non coding region depicted in nucleotides 51039-41739 of SEQ ID NO:4, a 3'-non-coding region depicted in nucleotides 9503-1 of SEQ ID NO:4, a contiguous intron-exon region between nucleotides 41738-9502 of SEQ ID NO:4, wherein a sequence segment comprising 41738-9502 of SEQ ID NO:4 encodes human mouse double minute 2 homolog depicted in SEQ ID NO:2, a	Page 10: Table 2

contiguous exon-intron region between nucleotide 41738-9502 of SEQ ID NO:4, wherein a sequence segment comprising 41738-9502 of SEQ ID NO:4 encodes human mouse double minute 2 homolog depicted in SEQ ID NO:2, an intron depicted in nucleotides 36385-40645, 36309-33127, 32994-29616, 29564-25577, 25507-25384, 25287-21169, 21006-14110, 13953-13267, and/or 13188-10665, a region comprising a dinucleotide of the following group: 41739-41738, 40645-40646, 36309-36310, 36384-36385, 32994-32995, 33126-33127, 29564-29565, 29615-29616, 25507-25508, 25287-25288, 25383-25384, 25576-25577, 21006-21007, 21168-21169, 14109-14110, 13953-13954, 13266-13267, 13188-13189, 10664-10665 and/or 9504-9503	
a transcription binding site selected from the group consisting of	
BINDING SITES	huMDM2, location in SEQ ID NO:4
AP1_C:	36-46, 2876-2886;
AP4_Q5:	7944-7980;
AP4_Q6:	7943-59, 8924-8940, 9294-9310;
ARNT_01:	1682-1706, 2193-2217, 9201-9225;
BRN2_01:	1040-1058, 7803-7821;
CAAT_01:	3292-3306;
CDPCR3HD_01:	6522-6540;
CEBPB_01:	1424-1438, 3917-3931, 4178-4192, 4787-4801, 6855-6869;
CREL_01:	5630-5642;
DELTAEF1_01:	83-95, 6328-6340;
FREAC7_01:	2757-2773, 5154-5170, 5823-5839;
GATA1_04:	4846-4858, 7017-7029;
GATA1_05:	8464-8476;
GATA2_02:	6045-6057, 6073-6085, 6142-6154;
GATA2_03:	2489-2501, 3323-3335, 3384-3396,

Page 9, line 29 to  
page 10, line 2;  
Table 3 on pages  
10-12



7393-7405:	
GATA3_02:	3264-3276, 6870-6882;
GATA3_03:	40-52, 5729-5741, 6529-6541, 6874-6886, 7041-7053, 7589-7601;
GATA_C: 7	349-7361, 8188-8200;
HFH2_01:	1743-1759, 7995-8011;
HFH3_01:	502-518, 1739-1755, 4160-4176, 9402-9418, 9418-9434;
HFH8_01:	8184-8200;
IK2_01:	951-963, 3588-3600;
MZF1_01:	1202-1210, 1447-1455, 4997-4005, 5424-5432;
NF1_Q6:	1480-1500, 8166-8182;
NFAT_Q6:	4190-4208, 6009-6027;
NKX25_01:	741-755, 1648-1662, 1885-1899, 1984-1998, 3609-3623, 4928-4942, 5060-5074, 5889-5903, 8850-8864, 9190-9204;
NKX25_02:	2584-2599, 2970-2984, 4644-4658, 5179-5193, 6482-6496;
NMYC_01:	2560-2572;
RORA1_01:	220-238, 2638-2656;
S8_01:	4644-4656, 4842-4854, 4845-4857, 5200-5212, 5371-5383, 5735-5747, 6482-6494, 6541-6553, 6544-6556, 6772-6784, 7270-7292, 7273-7285;
SOX5_01:	1355-1371, 1430-1446, 3094-3110, 3155-3171, 4669-4685, 4692-4708, 4789-4805;
SRV_02:	4164-4180, 5665-5681;
TATA_01:	1261-1277, 2574-2590, 2723-2739, 2733-

2749-2770-2786, 4199-4215, 4206-4222;	
TATA_C: 5900-5916, 7456-7472, 7702-7718, 7917-7933; and	
XFD2_01: 7702-7218, 7917-7933;	
a transcription binding site selected from the group consisting of	
BINDING SITES huMDM2, location in SEQ ID NO:4	
AP1_C: 12109-12119, 12695-12705, 22600-22610, 24166-24176, 31311-31321, 35234-35244, 39184-39194;	
AP1_Q2: 11952-11962, 12068-12078, 14798-14808, 21748-21758, 22613-22623, 23676-23686, 26562-26572, 30046-30056;	
AP1_Q4: 12695-12705, 31311-31321, 35234-35244, 36295-36305, 38784-38794, 39188-39198;	
AP4_Q6: 31635-31651;	
BRN2_01: 13448-13466, 14764-14782, 28094-28112, 40027-40045;	
CAAT_01: 11288-11302, 15054-15068;	
CDPCR3HD_01: 11286-11304, 13284-13302, 20846-20864, 29344-29362;	
CEBPB_01: 29241-29255;	
CREL_01: 36091-36103, 38873-38885;	
DELTAEF1_01: 18083-18095, 20385-20397, 26955-26967;	
FREAC7_01: 11982-11998, 15187-15202, 16523-16539, 16529-16545, 16587-16603, 16604-16620, 16676-16642, 16633-16649, 16644-16660, 16650-16666, 16657-16673, 16673-16689, 16762-16778, 21332-21348, 25689-25700, 26529-26545, 27767-27783, 29495-29511;	
GATA1_02: 10916-10928, 15775-15789, 18162-18174, 26088-26100, 32518-32530;	
GATA1_03: 28012-28024;	

GATA1_04:	11153-11165, 11630-11642, 13778-13790, 17439-17451, 19300-19312, 21606-21618, 22743-22755, 23747-23759, 25806-25818, 26529-26541, 29424-29436, 30455-30467, 32761-32778, 33352-33364, 33960-33972, 36101-36113, 40007-40019;
GATA1_05:	11590-11602, 26550-26562, 36737-36749;
GATA1_06:	18772-18784, 23054-23066, 35568-35580, 37855-37867;
GATA2_02:	20755-20767, 30830-30842, 34755-34767, 36285-36297, 39143-39155, 39641-39653, 40586-40598;
GATA2_03:	13535-13547, 22711-22723, 23161-23173, 25028-25040, 27237-27249, 36277-36289;
GATA3_02:	11558-11570, 16470-16482, 17225-17237, 19619-19631, 22156-22168, 22443-22455, 24713-24725, 27619-27631, 32716-32728, 34124-34136, 34163-34175, 36832-36844, 38403-38415;
GATA3_03:	10869-10881, 11515-11527, 13845-13857, 17221-17233, 18952-18964, 20050-20062, 40171-40183;
GATA_C:	15848-15860, 18899-18911, 23640-23652, 29072-29084, 30881-30893, 33198-33210, 37472-37484, 38621-38633;
GFI1_01:	35469-35481, 35492-35504;
HFH2_01:	15939-15955, 24636-24652, 25866-25882, 32171-32187, 35372-35388, 39457-35473;
HFH3_01:	13340-13356, 19218-19234, 21328-21344, 21336-21352, 21344-21360, 28062-28078, 32125-32141;
HFH8_01:	14133-14149, 22578-22584;
HNF3B_01:	13150-13166, 16505-16521, 25264-25280, 29443-29459, 37654-37670;
IK2_01:	11547-11559, 17144-17156, 18961-18973, 23883-23895, 27617-27629, 28908-28920, 29241-29253, 30752-30764, 34768-34780;

LYFL\_01: 12319-12331, 19191-19203, 37226-37238, 39430-39442;

MAX\_01: 22974-22986, 33339-33351;

MZF1\_01: 26105-26113, 35187-35195;

NF1\_Q6: 12048-12064, 33334-33354;

NFAT\_Q6: 13295-13313, 14157-14175, 14311-14329, 14414-14432, 18269-18287, 19326-19344, 20801-20819, 21177-21195, 22537-22555, 23861-23879, 25392-25410, 25879-25897, 27524-27542, 30636-30654, 30718-30736, 31525-31543, 33655-33673, 34726-34744, 34917-34935, 34990-35008, 35979-35997, 36479-36493, 36577-36595, 37154-37172, 40224-40242, 40365-40383;

NKX25\_01: 12041-12055, 12340-12354, 12471-12485, 12742-12756, 12877-12891, 13849-13863, 18995-19009, 21440-21454, 21883-21897, 28426-28440, 30964-30978, 32033-32047, 32265-32279;

NKX25\_02: 10998-11012, 12711-12725, 14131-14145, 14726-14740, 16024-16038;

NMYC\_01: 18753-18765, 18754-18766, 23076-23088, 30534-30546, 34400-34412;

RORA1\_01: 13134-13152, 22966-22984, 24934-24952, 33341-33359, 34760-34778;

S8\_01: 11000-11012, 11977-11989, 12048-12060, 12051-12063, 13747-13759, 13923-13935, 13926-13938, 14676-14688, 14679-14691, 16026-16038, 16313-16325, 16316-16328, 17515-17527, 20756-20768, 20759-20771, 23154-23166, 23157-23169, 25198-25210, 25201-25213, 26651-26663, 27508-27520, 27511-27523, 29450-29462, 29478-28490, 29775-29787, 29778-29790, 29813-29825, 29816-29828, 31329-31341, 31677-31689, 31680-31692, 31732-31744, 31735-31747, 36137-36149, 36140-36152, 36812-36824, 36815-36827, 37413-37425, 38679-38691, 39474-39486, 39477-39489;

SOX5\_01: 27397-27413, 27572-27588, 28100-28116, 29230-29246, 29439-29455, 30690-30706, 31595-31611, 33871-33887, 34113-34129, 34624-34640, 37668-37684, 38582-38598, 39124-

39140, 40410-40426;	
SRY_02: 20016-20032, 22410-22426, 27329-27345, 29162-29178, 29499-29515, 30646-30662, 31503-31519, 35928-35944, 37324-37340;	
TATA_01: 32722-32738, 32729-32745, 32807-32823, 33825-33841, 34120-34136, 35433-35449, 36593-36609;	
TATA_C: 11015-11031, 11817-11833, 13635-13651, 14930-14946;	
TCF11_01: 18543-18549, 22574-22580, 31281-31297, 31489-31505, 38754-38770;	
USF_01: 23075-23087, 32577-32589;	
VMYB_02: 11526-11538, 17384-17396, 18400-18412, 19549-19561, 22188-22200, 40486-40508 and	
XFD2_01: 16620-16636, 18153-18169, 22102-22118, 23141-23157.	
And a transcription binding site selected from the group consisting of	
BINDING SITES huMDM2 location in SEQ ID NO:4	
AP1_C: 44584-44594, 49069-49079;	
AP1_Q2: 42174-42184, 45217-45227, 48422-48422, 50447-50457;	
AP1_Q4: 42702-42712, 50806-50816;	
AP4_Q6: 42117-42133, 42118-42134, 42244-42260, 45432-45448; 45433-45449, 46609-46625;	
BRN2_01: 42310-42328, 44022-44040, 47514-47532, 48900-48918, 48967-48985;	
CAAT_01: 44866-44880;	
CDPCR3HD_01: 45671-45689, 49219-49237;	

CREL_01:	42437-42449, 49797-49809;	
FREAC7_01:	47026-47042, 47292-47308, 47658-47674;	
GATA1_02:	43482-43494, 48926-48938, 49284-49296;	
GATA1_03:	47371-47383;	
GATA1_04:	43054-43066, 43162-43162, 43967-43979, 45464-45476, 45916-45928, 47763-47775;	
GATA1_05:	49319-49331, 49459-49471;	
GATA1_06:	47590-47602;	
GATA2_02:	42660-42672, 43475-43487;	
GATA2_03:	43714-43726, 50948-50960;	
GATA3_02:	49155-49167, 49844-49856;	
GATA3_03:	42202-42214, 44810-44822, 48438-48450, 49136-49148, 49337-49349, 49869-49881;	
GATA_C:	44011-44023, 45256-45268, 45823-45835, 47915-47927, 49201-49213, 49573-49585;	
GFI1_01:	46606-46618, 47063-47075;	
HFH3_01:	47030-47046, 47284-47300, 47288-47304;	
IK2_01:	45275-45287;	
LYF1_01:	44564-44576, 46991-47003, 49567-49579;	
MAX_01:	43234-43246, 48726-48738;	
MZF1_01:	41772-41780, 42290-42298, 42295-42303, 44507-44515, 45105-45113, 45203-45211, 49948-49956, 50774-50782;	
NFI_Q6:	50209-50229;	
NEAT_Q6:	42061-42079, 44418-44436, 46399-46417,	

47974-47992, 49267-49285, 49964-49982, 50392-50410;	
NKX25_01:	42394-42408, 43507-43521, 46115-46129;
RORA1_01:	45073-45091, 48718-48736;
S8_01:	43552-43564, 45214-45226, 47160-47172, 48419-48431, 49295-49307, 50379-50391;
SOX5_01:	43716-43732, 46351-46367, 47156-47172, 47774-47790, 47868-47884, 47974-47990, 48915-48931, 50323-50339;
TATA_01:	45588-45604, 47625-47641, 48026-48042, 48659-48675, 49056-49072, 49079-49095, 49152-49168;
TCF11_01:	49115-49131;
VMYB_02:	42010-42022, 42279-42291, 44651-44663;
XFD2_01:	42870-42886, 42910-42926.

#### B. Independent Claim 24

Claim Element	Support in Specification
An isolated nucleic acid molecule 20-5000 contiguous nucleotides in length	Page 9, line 35 to page 10, line 2
consisting of a reverse or forward strand of a contiguous exon intron region between nucleotides 41738-9502 of SEQ ID NO:4 or contiguous intron-exon region between nucleotides 41738-9502 of SEQ ID NO:4,	Page 10, Table 2 and page 14, lines 29-32
wherein a sequence segment comprising 41738-9502 of SEQ ID NO:4 encodes human mouse double minute 2 homolog depicted in SEQ ID NO:2	Page 2, lines 7-13; page 10, Table 2

#### VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

- A. Whether claims 7, 10, 15-18, 20, 24, 25, 30 and 31 comply with the written description requirement under 35 USC §112, first paragraph
- B. Whether claim 7, 10, 15-18, 20, 24, 25, 30 and 31 are unpatentable over Muzny et al. , Genbank Accession No. AC025423 ("Muzny") in view of Vogelstein et al., US Patent No. 5,411,860 ("Vogelstein").

#### VII. ARGUMENT

##### A. Claims 7, 10, 15-18, 20, 24, 25, 30 and 31 Comply With the Written Description Requirement

In the final rejection dated January 2, 2009, it was asserted that claim 7, with dependent claims 10, 15-18, 20, 30 and 31, as amended on March 3, 2005 and August 29, 2005 and claim 24, with dependent claim 25, added on March 3, 2005 and amended on August 29, 2005 to recite "wherein a sequence segment comprising 41738-9502 of SEQ ID NO: 4 encodes human mouse double minute 2 homolog depicted in SEQ ID NO:2, ... a region comprising a dinucleotide of the following group: 41739-41738, ..... and/or 9504-9503" (lines 12-15 of claim 7) constitutes new matter and lacks support. Appellant had traversed and pointed out that there is more than adequate support in the specification for claims 7, 10, 15-18, 20, 30 and 31 and in particular, for the sequence segment 41738-9502, the **nucleotide ranges** for the various sites for huMDM2 location in SEQ ID NO:4, specific ranges that relate to the specific dinucleotide ranges or specific 'exon/intron' organization in terms of the



specific nucleotide range(s) and in particular, the region comprising a dinucleotide i.e., nucleotides '41739-41738'.

In response, the Advisory Action on page 3 states

Applicant does not indicate and the examiner is unable to locate adequate support in the specification for such positions/ranges in SEQ 10 NO: 4. It is also noted that claim 7, starting from line 17, indicate *nucleotide ranges* for various binding sites for huMDM2, location in SEQ 10 NO:4 (claims filed 9/14/2008), pages 2-8. However, there is no basis for these nucleotide ranges (For Example: AP1\_C: 36-46, 2876-2886; AP4\_Q5: 7944-7980 .....

Thus there is no indication that the specific segments or ranges were within the scope of the invention as conceived by Applicants at the time the application was filed.

The Advisory Action on page 5 further states

Applicants arguments are considered but not found to be persuasive because the specification as originally filed does not teach the specific ranges that related to the specific dinucleotide ranges or the specific '*exon/intron*' organization in term of the specific nucleotide range(s).

Based upon the teachings of Table 2, on page 10, wherein Exon 1 begins at nucleotide 40726 of SEQ ID NO:4 and the stop codon terminates at nucleotide 10091 of SEQ ID NO:4. However is it not clear that nucleotides 41738-9502 of SEQ ID NO: 4 would constitute a sequence segment that encodes human mouse double minute 2 homolog protein. Similarly, the region comprising a dinucleotide i.e., nucleotides '41739-41738', does not have no basis in the instant specification.

In response, and as stated in previous responses, there is more than adequate support in the specification for the positions/ranges in SEQ ID NO:4 recited in claims 7, 24 and 25.

On page 4, lines 25-30 of the specification, it is stated:

The invention is directed to isolated genomic polynucleotide fragments that encode ...human mouse double minute 2 homolog, which in a specific embodiment are...human mouse double minute 2 homolog genes, as well as vectors and hosts containing these fragments and polynucleotide fragments hybridizing to noncoding regions, as well as antisense oligonucleotides to these fragments.

Further, the specification on page 14, lines 29-33 states:

The present invention also relates to nucleic acid constructs comprising a polynucleotide sequence containing the exon/intron segments of the .....human mouse double minute 2 homolog gene (nucleotides 1-51039 of SEQ ID NO:4).

Table 2 on page 10 of the specification shows "exon/intron organization of the human mouse double minute 2 homolog gene... in SEQ ID NO:4, 51039 base pairs; nucleotides 99541-150579 in the genomic clone of accession no. AC025423 (reverse strand cloning)". Exon 1 according to Table 2 begins at nucleotide 40726 of SEQ ID NO:4 and the stop codon terminates at nucleotide 10091 of SEQ ID NO:4. However, there are further nucleotide sequences set forth in SEQ ID NO:4, nucleotides 51039-40727 at the 5' end and nucleotides 1-10090 at the 3' end. It is merely implicit in Table 2 that the minimum genomic DNA sequence encoding human mouse double minute 2 homolog ("MDM2") protein ranges from nucleotides 40726-10091. Certainly a larger sequence can encode human MDM2 protein, such as the recited nucleotides 41738-9502 of SEQ ID NO:4. As noted above, the human MDM2 gene is actually encompassed by nucleotides 1-51039 of SEQ ID NO:4). Thus, there is support for the recitation in claims 7 for the phrase "wherein a sequence segment comprising 41738-9502 of SEQ ID NO: 4 encodes human MDM2 depicted in SEQ ID NO:2, ... a region comprising a dinucleotide of the following group: 41739-41738, ..... and/or 9504-9503" (lines 12-15 of claim 7).

Further, there is support for the recitation "a region comprising a dinucleotide of the following group: 41739-41738" in claims 7 and 25. In particular, Applicant notes that page 9, lines 29-32 states:

The invention is further directed to polynucleotide fragments containing or hybridizing to noncoding regions of the ...human mouse double minute 2 homolog genes. These include but are not limited to an expression control element, an intron, a 5'-non-coding region, a 3'-non-coding region and splice junctions (see tables 1-2, as well as transcription factor binding sites (see Table 3).

A region encompassing the dinucleotide 41739-41738 would be within the 3'-noncoding region and would thus constitute a fragment containing a 3' non-coding region.

Similarly, it is Appellant's position that the dinucleotide ranges of the specific exon/intron organization would be taught as well. This is because the various exon/intron (splice) junctions are disclosed in Table 2 and it is clearly indicated that fragments containing

these exon/intron junction would be encompassed as well. Specifically it is stated on page 9, line 33 to page 10, line 44:

The polynucleotide fragments may be a short polynucleotide fragment that is between about 20 nucleotides to about 50 nucleotides in length. Such shorter fragments may be useful for diagnostic purposes. Such short polynucleotide fragments are also preferred with respect to polynucleotides containing or hybridizing to polynucleotides containing splice junctions. Alternatively larger fragments, e.g., of about 50, 150, 500, 600, 2000 or about 5000 nucleotides in length may be used.

Finally, with respect to lack of support of transcription binding sites recited in claim 7, it is Appellant's position that Table 2 lists various transcription factor binding sites. Each site is readily found in SEQ ID NO:4. Appellant notes that fragments 20 nucleotides in length are disclosed in the specification as well (see p. 10, lines 13-18). It is stated in the specification:

...These include but are not limited to an intron, a 5'-non-coding region, a 3'-noncoding region and splice junctions (see table 1), as well as transcription factor binding sites (see table 2). The polynucleotide fragments may be a short polynucleotide fragment which is at least 15, 16, 17, 18 or 19 nucleotides in length but may be between about 20 nucleotides to about 50 nucleotides in length. Such shorter fragments may be useful for diagnostic purposes.

One of ordinary skill in the art would given the information provided in Table 2 and tools known to one of ordinary skill in the art as of the priority date been able to identify the transcription sites recited in claim 7 such as the program MatInspector and is commercially available from Geomatix GmbH ([www.geomatix.de](http://www.geomatix.de)). Such programs were available as of the priority date of the instant application. Furthermore, as noted in MPEP 2163,

While there is no *in haec verba* requirement, newly added claim limitations must be supported in the specification through express, implicit, or inherent disclosure.

Given that there is information in Table 2 indicating the presence of all of the recited transcription factor binding sites, given that the sequences of these transcription factor binding sites are well known in the art, given that one of ordinary skill in the art could have identified the transcription binding sites recited and given that SEQ ID NO:4 is known, there is sufficient support in the specification to meet the criteria set forth in the MPEP. Thus, no new matter has been added with respect to claim 7.

In conclusion, there is adequate support for the phrase in claim 7 "wherein a sequence segment comprising 41738-9502 of SEQ ID NO:4 encodes human mouse double minute 2 homolog depicted in SEQ ID NO:2, ... a region comprising a dinucleotide of the following group: 41739-41738, ... " as well as claim 24 directed to an isolated nucleic acid molecule 20-5000 nucleotides in length consisting of a reverse or forward strand of a contiguous exon-intron region or intron-exon region between nucleotides 41738-9502 of SEQ ID NO:4 and claim 25 directed to an isolated nucleic acid molecule 20-5000 nucleotides in length comprising nucleotides 41739-41738....". Further, claims 10, 15-18, 20, 24, 25, 30 and 31 ultimately depend from claim 7. Thus, arguments made with respect to claim 7 would apply to these claims as well.

**B. Claims 7, 10, 15-18, 20, 24, 25, 30 and 31 are not Obvious over Muzny et al. in view of Vogelstein**

In the final rejection dated January 2, 2009, the Office Action stated:

It would have been obvious to one of ordinary skill in the art at the time the invention was made to use said cDNA to identify the genomic DNA that encodes the human MDM2 homolog of SEQ ID NO:2. The motivation is provided by Vogelstein et al. who teach that it binds to oncogene p53 and is diagnostic of tumorigenesis (citation omitted). The state of the art provides various techniques for obtaining genomic DNA using cDNA probes that are usually labeled. The comparison of genomic and cDNA would result in the identification of regions comprising exon-intron and intron-exon junctions within coding and noncoding regions. One of ordinary skill in the art would have been motivated to use said non-coding regions or fragments thereof of at least 20 nucleotides and up to 5000 or 51039 nucleotides (the entire length of SEQ ID NO:4) nucleotides for detecting splice variants of the genomic DNA encoding human MDM2 homolog in genomic nucleotide samples from an individual, for example. As a matter of convenience a non-coding region such as an exon-intron or intron-exon region or fragments thereof can be present in a kit or on a solid support. Further, said support can be a microarray according to a customary use of nucleic acid molecules in the art.

Appellant had traversed the rejection. In the response dated May 30, 2009, Appellant had asserted that it would not have been obvious to combine the disclosure of Muzny and

Vogelstein given that there was no suggestion to do so and further to assert that undue experimentation would have been required not only to locate the MDM2 gene but also identify exon-intron junctions.

In response, in the Advisory Action dated June 24, 2009, the following assertions are made:

With regard to the 103(a) rejection, Applicant's arguments can be summarized as follows: 1) Muzny contains a small portion of chromosome 12, where the location of the MDM2 gene was indicated by Vogelstein as 12q12-14 while Applicant found it on 12q and 2) the MDM cDNA constitutes only 1.6% of the clone disclosed by Muzny, therefore it would be undue experimentation to locate the MDM2 gene and identify its exon-intron junctions.

Applicant further argues Second, Appellant asserts that there would not be a reasonable expectation of success of obtaining the claimed non-coding sequences of SEQ ID NO: 4 in view of the cited references. Vogelstein placed the human MDM2 homologue gene at 12q12-14. As noted above, there was actually a previous disclosure stating that the MDM2 was located between 12q14.3-15 (see, for example, Andersen et al., 1996, *Mammalian Genome* 7: 780-783 and Bureau, 1995, *Genomics* 28: 109-112, submitted and disclosed in previous response attached hereto as Exhibit 1). However, given the conflicting locations published as of the priority, one of ordinary skill in the art would not have known which location was actually correct" (Brief, page 14).

This is not agreed with because the actual location does not matter as long as it is a part of the Muzny sequence, which it is. Applicant did not need to separate the Muzny sequence into the fragments containing different arms of chromosome 12. In fact, Applicant did not isolate the fragment 12q12-14 or 12q14.3-15. He run cDNA against the genomic DNA disclosed by Muzny and found the location of the gene where it was. This experiment was performed according to the knowledge and the state of the art as evidenced by Watson et al. Watson et al. teach that "once the first genes were cloned, introns were identified by comparing the cloned genomic DNA with the corresponding cloned cDNA" ("Recombinant DNA", page 137, 2<sup>nd</sup> column, form PFO892 mailed 4/16/07). Applicant's argument would be convincing if the exact location would need to be known before the comparison of the genomic and the cDNA is made. This is not the case because the work is done on the genomic DNA that is known without fragmentation thereof. Applicants further

argues that "Watson would not apply in this case since in Watson, the genes themselves were actually cloned" (*ibid.*, page 18, last sentence). This is not persuasive because Muzny provided the piece of the genomic DNA containing the requisite gene. Having the cDNA, it does not require undue experimentation to identify the fragment of the genomic DNA corresponding to the gene and exon-intron locations within said gene.

The second type of Applicant's arguments concerns with the fact that the cDNA constitutes only 1.6% of the genomic DNA. While a large quantity of the experimentation may be involved, it is not undue because sufficient guidance and knowledge are provided by the art.

Appellant maintains that the claimed invention is not obvious over the cited references for several reasons. (1) One of ordinary skill in the art could not have predicted where the human MDM2 gene was located; (2) The conflicting teachings in the art with respect to the possible location of the human MDM2 gene is of particular significance; (3) At best, this would still be an "obvious to try" situation; (4) There was a long felt need and it is likely that others had tried but failed before the priority date of the instant application. Each of these reasons will be discussed in further detail below:

**(1) One of ordinary skill in the art could not have predicted where the MDM2 gene was located**

Muzny merely discloses the sequence of a genomic clone containing chromosome 12 sequences, not isolated SEQ ID NO:4 (see Evidence Appendix). No indication is provided in the Muzny disclosure as to which portion of chromosome 12 has actually been sequenced. Vogelstein (see Evidence Appendix) merely discloses the human MDM2 cDNA. Appellant was the first to identify SEQ ID NO:4 and determine that it did indeed encode human MDM2 and in particular determine the claimed noncoding sequences.

Muzny only disclosed the sequence of the clone AC025423. There was no indication at the time of the filing date of the instant application that AC025423 actually contained SEQ ID NO:4, more specifically a sequence encoding human MDM2. Muzny only sequenced a small portion of chromosome 12, the chromosome 12 BAC clone AC025423 which contains less than 0.12% of the chromosome (about 157 kB vs. about 132,000 Kb). The clone's sub-

site on chromosome 12 apparently was not then known. Muzny did not even suggest, much less state whether the clone AC025423 did or did not contain any gene. The determination of a sequence of genomic DNA does not necessarily imply the presence of a gene. This is because it is well known in the art and was certainly well known when the instant invention was made and as of the filing date of the instant application that most genomic DNA contains "junk" DNA, not genes (see, e.g., Wong et al., 2000, "Is Junk DNA Mostly Intron DNA?" Genome Research 10: 1672-1678 in Evidence Appendix that discuss "junk" DNA in further detail). As is evident from this review article, junk DNA was well known as of the priority date of the above-referenced application. When obtaining genomic clone sequences, one of skill in the art would not have any way of knowing whether or not it actually contains a gene(s) or just junk DNA. Even noting that a clone has a high GC content is a poor guide to a sequence's likely gene content, especially given the presence of pseudogenes. Therefore, *contra* to assertions made in the Advisory Action, disclosure of this particular sequence would not indicate to the skilled artisan that this clone would necessarily contain a gene.

The clone AC025423 is 150,579 nucleotides in length. The cDNA sequence only contains 2372 nucleotides (1.6% of AC025423). However, one of ordinary skill in the art would not know where or how these 2372 nucleotides are interspersed within the AC025423 clone or if indeed it is even present. No teachings are provided as to the structure of the MDM2 gene itself: number and size of exons, number and size of introns, locations of exons and introns and number and size of 5' and 3' untranslated regions. The possibilities are close to infinite. Thus there would not be a reasonable expectation of success of obtaining the claimed sequences given that the isolation and identification of the claimed sequences constitutes undue experimentation.

Appellant wishes to further emphasize that the size of the cDNA relative to the isolated genomic clone is of particular significance. If, for example, the cDNA constituted 50% of a particular genomic clone, considerable less experimentation would be involved in determining the sequence of a particular gene than when it is merely 1.6% of the genomic clone.

Watson has been mentioned during prosecution to support the premise that it would be simple to locate intron/exon junctions in the Muzny sequence. Appellant asserts that the teachings of Watson would be of little significance. Watson et al. provides a general

description as to how introns may be located by comparing cloned cDNA to cloned genomic DNA. However, it would not be feasible to compare cDNA to the entire Muzny sequence. Smaller fragments would be necessary. As will be discussed in further detail below, trying to find the MDM2 gene given the teachings of Muzny, Vogelstein and even Watson would metaphorically be like “throwing darts”.

**(2) The conflicting teachings in the art with respect to the possible location of the MDM2 gene is of particular significance**

Appellants disagree with the assertion that the “actual location does not matter as long as it is a part of the Muzny sequence, which it is”. As noted in previous responses, Vogelstein placed the human MDM2 homologue gene at 12q12-14. Actually, this finding is incorrect. After the publication of Vogelstein, the gene was found not to be located at 12q12-14. The gene is actually several millions of base pairs away at 12q15 (see, for example, Andersen et al., 1996, *Mammalian Genome* 7:780-783, Bureau, 1995, *Genomics* 28: 109-112 and Genecard and set forth in the Evidence Appendix, attached hereto as Exhibit 1 and previously made of record). Thus, even if these two references were indeed combined, the ordinary skilled artisan would have looked for the MDM2 gene in the wrong location and thus would not have obtained the claimed sequences. Once one of ordinary skill in the art was unable to find the human MDM2 gene at 12q12-14, one of ordinary skill would have had no clear direction as to where the MDM2 gene was located on that particular sequence or if indeed it would be located on that sequence at all. The human MDM2 gene is interspersed among the recited Muzny sequence. Thus the possible location of the MDM2 gene is of particular significance since it would be less likely that one of ordinary skill in the art would have a reasonable expectation of success.

**(3) At best, this would still be an “obvious to try” situation**

It is Appellant’s view that the combination of the two references would still constitute “obvious to try” even under *KSR v. Teleflex*. *KSR v. Teleflex*, 82 USPQ2d 1385; 127 S.Ct. 1727 (2007) and *In re Kubin*, 561 F.3d 1351 (Fed. Cir. 2009). Appellant notes that the “Examination Guidelines for Determining Obviousness Under 35 U.S.C. 103 in View of the Supreme Court Decision in *KSR International Co. v. Teleflex*”, 72 FR 57526 (October 10,



2007) (hereinafter “Examination Guidelines”) stated with respect to “obvious to try” the following:

To reject a claim based on this rationale, Office personnel must resolve the *Graham* factual inquiries. Office personnel must then articulate the following:

(1) a finding that at the time of the invention, there had been a recognized problem or need in the art, which may include a design need or market pressure to solve a problem;

(2) a finding that there had been a finite number of identified, predictable potential solutions to the recognized need or problem;

(3) a finding that one of ordinary skill in the art could have pursued the known potential solutions with a reasonable expectation of success; and

(4) whatever additional findings based on the *Graham* factual inquiries may be necessary, in view of the facts of the case under consideration, to explain a conclusion of obviousness.

*Ex parte Kubin*, 83 USPQ2d 1410, 2007 WL 2070495 (Bd. App. & Int. 2007) was cited as an example in the Examination Guidelines of a situation where a finite number of identified, predictable solutions are provided with a reasonable expectation of success. However, in the Federal Circuit decision affirming *Ex parte Kubin*, *In re Kubin* 561 F.3d 1351 (Fed. Cir. 2009), the Court stated:

To differentiate between proper and improper applications of “obvious to try,” this court outlined two classes of situations where “obvious to try” is erroneously equated with obviousness under § 103. In the first class of cases, what would have been “obvious to try” would have been to vary all parameters or try each of numerous possible choices until one possibly arrived at a successful result, where the prior art gave either no indication of which parameters were critical or no direction as to which of many possible choices is likely to be successful (citing *In re O’Farrell*, 853 F.2d 894, 903 (Fed. Cir. 1988)). In such circumstances, where a defendant merely throws metaphorical darts at a board filled with combinatorial prior art possibilities, courts should not succumb to hindsight claims of obviousness. The inverse of this proposition is succinctly encapsulated by the Supreme Court’s statement in *KSR* that where a skilled artisan merely pursues “known options” from a “finite number of

identified, predictable solutions,” obviousness under § 103 arises. 550 U.S. at 421.

The second class of *O’Farrell’s* impermissible “obvious to try” situations occurs where what was “obvious to try” was to explore a new technology or general approach that seemed to be a promising field of experimentation, where the prior art gave only general guidance as to the particular form of the claimed invention or how to achieve it.

As stated above, the teaching of the prior art would constitute metaphorically throwing the darts. Muzny et al. merely discloses a chromosome 12 genomic sequence. No direction is provided in Muzny regarding the location of the MDM2 gene. Further, The clone AC025423 is 150,579 nucleotides in length. No indication is even given as to where the MDM2 gene is contained within this clone. The cDNA sequence only contains 2372 nucleotides (1.6% of AC025423). However, one of ordinary skill in the art would not know where or how these 2372 nucleotides are interspersed within the AC025423 clone or if indeed it is even present. No teachings are provided as to the structure of the MDM2 gene itself: number and size of exons, number and size of introns, locations of exons and introns and number and size of 5’ and 3’ untranslated regions. Additionally, the situation is further complicated by the two disclosures in the art regarding the possible location of the Muzny sequence, Vogelstein and Andersen. Specifically as noted above, Appellant notes that Vogelstein placed the human MDM2 homolog gene at 12q12-14. After publication of Vogelstein, the MDM2 homolog gene was found to be located several kilobases away at 12q15. One of skill in the art would not have known where to look.

The second “obvious to try” situation would apply as well. This is because only general guidance is provided in the art with respect to detecting intron/exon junctions. Specifically, Watson et al. on pages 137-138, states

Once the first genes were cloned, introns were identified by comparing the cloned genomic DNA with the corresponding cloned cDNA. For small genes, such as the beta-globin gene, the sizes of the introns and the locations of intron-exon boundaries were precisely determined by sequencing cloned genomic DNA and comparing the sequence with the cDNA sequence and with the protein sequence. Introns exist in genes from all eukaryotic animals, in plant genes, and, surprisingly, in genes of the *E. coli* phage T4. Often the introns of a gene contain many more nucleotides than do its coding exons. thus

accounting for the previously unexplained large sizes of many primary RNA transcripts. The number and size of introns vary widely from one gene to another.

There is no specific methodology provided regarding how to locate the human MDM2 gene itself or even its size. Further, even Watson admits to variability of size and location of introns.

Further, the claimed invention can be distinguished from the facts in *Kubin*.

Specifically, in *Kubin*, the claimed invention was directed to

An isolated nucleic acid molecule comprising a polynucleotide encoding a polypeptide at least 80% identical to amino acids 22-221 of SEQ ID NO:2, wherein the polypeptide binds CD48.

*Id.*

In the rejection affirmed by the Board, the Examiner asserted that

The skilled artisan would have been motivated to isolate the nucleic acid sequence corresponding to NAIL, based on Valiante's disclosure of p38 (which is the same protein as NAIL) and Valiante's express teachings how to isolate p38 cDNA by using conventional techniques, such as taught in Sambrook, including using mAbC1.7, a probe specific for p38.

*Id.*

This finding of obviousness in *Kubin* appeared to be predicated on the obviousness of isolating NAIL cDNA, not any other nucleic acid sequences encoding the NAIL protein. *In re Kubin* specifically states

Moreover, the record strongly reinforces ...the board's factual finding that one of ordinary skill would have been motivated to isolate NAIL cDNA, given Valiante's teaching that p38 is "Expressed by virtually all human NK cells and thus plays a role in the immune response"....The record shows that the prior art teaches a protein of interest, a motivation to isolate the gene coding for that protein, and illustrative instructions to use a monoclonal antibody specific to the protein for cloning this gene.

*Id.*

In contrast, the claimed invention is directed to a nucleic acid molecule 20-51039 contiguous nucleotides in length consisting of a reverse or forward strand of a region of SEQ ID NO:4. SEQ ID NO:4 is a genomic sequence, not cDNA encoding MDM2. This is very different from *Kubin* which involves deducing the NAIL cDNA sequence from NAIL.

polypeptide sequence and subsequently isolating this sequence. The art cited in the instant application was a cDNA sequence and a genomic clone containing chromosome 12 sequence. Considerably more and undue experimentation would be involved in identifying and isolating the genomic sequence and consequently the claimed regions comprising the recited noncoding regions of the MDM2 gene than in identifying cDNA in view of a disclosure of a polypeptide sequence. The claimed sequence is between 20-51039 nucleotides in length. The Muzny sequence is 150,579 nucleotides in length. No direction or frame of reference is provided as to where the MDM2 sequence could be located. Furthermore, although Vogelstein, disclosed the MDM2 cDNA, no direction is provided as to the location of exons, introns, 5' untranslated and/or 3' untranslated regions.

It is Appellant's view that the claimed invention is more analogous to *Ex Parte Koide*, Appeal 2009-1912, 2009 WL 1719400 (Bd. Pat. App. & Interf. 2009) decided subsequent to *In re Kubin*. In *Koide*, it was held that claims to a fibronectin molecule containing a stabilizing mutation of at least one of Asp7, Asp23 or Glu9 were not obvious over a prior art disclosure of randomly screening Fibronectin type III polypeptides at a variety of positions and the disclosure of a Fibronectin type III subunit binding domain since "there is no predictable expectation that performing the random screening methods of the cited prior art would predictably result in the polypeptides containing the recited mutations. Here, there is no indication that performing the recombinant DNA methods would necessarily result in the claimed nucleic acid molecules.

**(4) There was a long felt need and it is likely that others had tried but failed before the priority date of the instant application.**

There has been a great deal of interest in the scientific community in human MDM2 given its potential use as a diagnostic and therapeutic agent. This interest is summarized in Vogelstein. Additionally, Appellant submitted during prosecution an IDS listing four references relating to MDM2 (a copy of the 1449 Form previously submitted and made of record and included in Evidence Appendix). However, there was absolutely no disclosure or suggestion of the genomic organization of MDM2 genomic DNA until the instant application was filed. An independent disclosure of the genomic organization of the human MDM2 gene was not available until July 21, 2004, more than one year after the filing date of the instant

application (see Liang et al., 2004, Gene 338:217-223, previously made of record in the Supplemental Appeal Brief submitted March 2, 2008). The dearth of knowledge regarding the human MDM2 gene was actually admitted to by Liang et al. on the first page of his article where it was stated "Although the human MDM2 cDNA sequence has been reported, the genomic organization of the human gene has not been documented". Further, Liang et al. discussed the usefulness of determining the genomic structure and organization of the human MDM2 gene on page 218:

One of the distinctive properties of MDM2 is the possession of an extremely complex expression pattern. Its multiple-sized transcripts and proteins have been found in tumour samples and cell lines by a number of groups (citations omitted). In our previous studies, five alternatively sized transcripts of the human MDM2 were found in human ovarian tumour, bladder tumour and leukaemic cell samples (citation omitted). ...Here we present data demonstrating two further MDM2 transcript forms with internal sequence deletions in human tumour tissue. We Hypothesised that these transcripts are generated by alternative splicing. To test this hypothesis and to explore the associated mechanisms, we have investigated the genomic structure and organization of the human MDM2 gene...

It is well established case law that secondary considerations such as commercial success, long felt need, unexpected results must be considered before making an obviousness determinations (see, for example, *Shackelton v J Kaufman Iron Works, Inc.*, 689 F.2d 334, (Fed Cir 1982) and *Ortloff v Gulsby Engineering*, 884 F2d 1399, 2 USPQ2d 1335 (Fed. Cir. 1989), reh'g denied Sept. 1989).

Respectfully submitted,

/Cheryl H Agris/

Date: October 2, 2009

Cheryl H. Agris, Reg. No. 34,086  
P.O. Box 8459  
Pelham, N.Y. 10803  
(914) 712-0093

## CLAIMS APPENDIX

7. An isolated nucleic acid molecule 20-51039 contiguous nucleotides in length consisting of a reverse or forward strand of a region of SEQ ID NO:4, wherein said region is selected from the group consisting of a 5'-non coding region depicted in nucleotides 51039-41739 of SEQ ID NO:4, a 3'-non-coding region depicted in nucleotides 9503-1 of SEQ ID NO:4, a contiguous intron-exon region between nucleotides 41738-9502 of SEQ ID NO:4, wherein a sequence segment comprising 41738-9502 of SEQ ID NO:4 encodes human mouse double minute 2 homolog depicted in SEQ ID NO:2, a contiguous exon-intron region between nucleotide 41738-9502 of SEQ ID NO:4, wherein a sequence segment comprising 41738-9502 of SEQ ID NO:4 encodes human mouse double minute 2 homolog depicted in SEQ ID NO:2, an intron depicted in nucleotides 36385-40645, 36309-33127, 32994-29616, 29564-25577, 25507-25384, 25287-21169, 21006-14110, 13953-13267, and/or 13188-10665, a region comprising a dinucleotide of the following group: 41739-41738, 40645-40646, 36309-36310, 36384-36385, 32994-32995, 33126-33127, 29564-29565, 29615-29616, 25507-25508, 25287-25288, 25383-25384, 25576-25577, 21006-21007, 21168-21169, 14109-14110, 13953-13954, 13266-13267, 13188-13189, 10664-10665 and/or 9504-9503; a transcription binding site selected from the group consisting of

BINDING SITES                      huMDM2, location in SEQ ID NO:4

AP1\_C:                                36-46, 2876-2886;

AP4\_Q5:                              7944-7980;

AP4_Q6:	7943-59, 8924-8940, 9294-9310;
ARNT_01:	1682-1706, 2193-2217, 9201-9225;
BRN2_01:	1040-1058, 7803-7821;
CAAT_01:	3292-3306;
CDPCR3HD_01:	6522-6540;
CEBPB_01:	1424-1438, 3917-3931, 4178-4192, 4787-4801, 6855-6869;
CREL_01:	5630-5642;
DELTAEFI_01:	83-95, 6328-6340;
FREAC7_01:	2757-2773, 5154-5170, 5823-5839;
GATA1_04:	4846-4858, 7017-7029;
GATA1_05:	8464-8476;
GATA2_02:	6045-6057, 6073-6085, 6142-6154;
GATA2_03:	2489-2501, 3323-3335, 3384-3396, 7393-7405;

GATA3_02:	3264-3276, 6870-6882;
GATA3_03:	40-52, 5729-5741, 6529-6541, 6874-6886, 7041-7053, 7589-7601;
GATA_C: 7	349-7361, 8188-8200;
HFH2_01:	1743-1759, 7995-8011;
HFH3_01:	502-518, 1739-1755, 4160-4176, 9402-9418, 9418-9434;
HFH8_01:	8184-8200;
IK2_01:	951-963, 3588-3600;
MZF1_01:	1202-1210, 1447-1455, 4997-4003, 5424-5432;
NF1_Q6:	1480-1500, 8166-8182;
NFAT_Q6:	4190-4208, 6009-6027;
NKX25_01:	741-755, 1648-1662, 1885-1899, 1984-1998, 3609-3623, 4928-4942, 5060-5074, 5889-5903, 8850-8864, 9190-9204;
NKX25_02:	2584-2599, 2970-2984, 4644-4658, 5179-5193, 6482-6496;



NMYC\_\_01: 2560-2572;

RORA1\_\_01: 220-238, 2638-2656;

S8\_\_01: 4644-4656, 4842-4854, 4845-4857, 5200-5212, 5371-5383, 5735-5747,  
6482-6494, 6541-6553, 6544-6556, 6772-6784, 7270-7292, 7273-7285;

SOX5\_\_01: 1355-1371, 1430-1446, 3094-3110, 3155-3171, 4669-4685, 4692-4708,  
4789-4805;

SRY\_\_02: 4164-4180, 5665-5681;

TATA\_\_01: 1261-1277, 2574-2590, 2723-2739, 2733-2749, 2770-2786, 4199-4215,  
4206-4222;

TATA\_\_C: 5900-5916, 7456-7472, 7702-7718, 7917-7933; and

XFD2\_\_01: 7702-7218, 7917-7933; .

a transcription binding site selected from the group consisting of

BINDING SITES      huMDM2, location in SEQ ID NO:4

API\_\_C: 12109-12119, 12695-12705, 22600-22610, 24166-24176, 31311-31321, 35234-  
35244, 39184-39194;

AP1\_Q2: 11952-11962, 12068-12078, 14798-14808, 21748-21758, 22613-22623, 23676-23686, 26562-26572, 30046-30056;

AP1\_Q4: 12695-12705, 31311-31321, 35234-35244, 36295-36305, 38784-38794, 39188-39198;

AP4\_Q6: 31635-31651;

BRN2\_01: 13448-13466, 14764-14782, 28094-28112, 40027-40045;

CAAT\_01: 11288-11302, 15054-15068;

CDPCR3HD\_01: 11286-11304, 13284-13302, 20846-20864, 29344-29362;

CEBPB\_01: 29241-29255;

CREL\_01: 36091-36103, 38873-38885;

DELTAEF1\_01: 18083-18095, 20385-20397, 26955-26967;

FREAC7\_01: 11982-11998, 15187-15202, 16523-16539, 16529-16545, 16587-16603, 16604-16620, 16676-16642, 16633-16649, 16644-16660, 16650-16666, 16657-16673, 16673-16689, 16762-16778, 21332-21348, 25689-25700, 26529-26545, 27767-27783, 29495-29511;

GATA1\_02: 10916-10928, 15775-15789, 18162-18174, 26088-26100, 32518-32530;

GATA1\_03: 28012-28024;

GATA1\_04: 11153-11165, 11630-11642, 13778-13790, 17439-17451, 19300-19312, 21606-21618, 22743-22755, 23747-23759, 25806-25818, 26529-26541, 29424-29436, 30455-30467, 32761-32778, 33352-33364, 33960-33972, 36101-36113, 40007-40019;

GATA1\_05: 11590-11602, 26550-26562, 36737-36749;

GATA1\_06: 18772-18784, 23054-23066, 35568-35580, 37855-37867;

GATA2\_02: 20755-20767, 30830-30842, 34755-34767, 36285-36297, 39143-39155, 39641-39653, 40586-40598;

GATA2\_03: 13535-13547, 22711-22723, 23161-23173, 25028-25040, 27237-27249, 36277-36289;

GATA3\_02: 11558-11570, 16470-16482, 17225-17237, 19619- 19631, 22156-22168, 22443-22455, 24713-24725, 27619-27631, 32716-32728, 34124-34136, 34163-34175, 36832-36844, 38403-38415;

GATA3\_03: 10869-10881, 11515-11527, 13845-13857, 17221-17233, 18952-18964, 20050-20062, 40171-40183;

GATA\_C: 15848-15860, 18899-18911, 23640-23652, 29072-29084, 30881-30893, 33198-33210, 37472-37484, 38621-38633;

GPII\_01: 35469-35481, 35492-35504;

HFH2\_01: 15939-15955, 24636-24652, 25866-25882, 32171-32187, 35372-35388, 39457-35473;

HFH3\_01: 13340-13356, 19218-19234, 21328-21344, 21336-21352, 21344-21360, 28062-28078, 32125-32141;

HFH8\_01: 14133-14149, 22578-22584;

HNFB3B\_01: 13150-13166, 16505-16521, 25264-25280, 29443-29459, 37654-37670;

IK2\_01: 11547-11559, 17144-17156, 18961-18973, 23883-23895, 27617-27629, 28908-28920, 29241-29253, 30752-30764, 34768-34780;

LYF1\_01: 12319-12331, 19191-19203, 37226-37238, 39430-39442;

MAX\_01: 22974-22986, 33339-33351;

MZP1\_01: 26105-26113, 35187-35195;

NF1\_Q6: 12048-12064, 33334-33354;

NFAT\_Q6: 13295-13313, 14157-14175, 14311-14329, 14414-14432, 18269-18287, 19326-19344, 20801-20819, 21177-21195, 22537-22555, 23861-23879, 25392-25410, 25879-25897, 27524-

27542, 30636-30654, 30718-30736, 31525-31543, 33655-33673, 34726-34744, 34917-34935, 34990-35008, 35979-35997, 36479-36493, 36577-36595, 37154-37172, 40224-40242, 40365-40383;

NKX25\_01: 12041-12055, 12340-12354, 12471-12485, 12742-12756, 12877-12891, 13849-13863, 18995-19009, 21440-21454, 21883-21897, 28426-28440, 30964-30978, 32033-32047, 32265-32279;

NKX25\_02: 10998-11012, 12711-12725, 14131-14145, 14726-14740, 16024-16038;

NMYC\_01: 18753-18765, 18754-18766, 23076-23088, 30534-30546, 34400-34412;

RORA1\_01: 13134-13152, 22966-22984, 24934-24952, 33341-33359, 34760-34778;

S8\_01: 11000-11012, 11977-11989, 12048-12060, 12051-12063, 13747-13759, 13923-13935, 13926-13938, 14676-14688, 14679-14691, 16026-16038, 16313-16325, 16316-16328, 17515-17527, 20756-20768, 20759-20771, 23154-23166, 23157-23169, 25198-25210, 25201-25213, 26651-26663, 27508-27520, 27511-27523, 29450-29462, 29478-28490, 29775-29787, 29778-29790, 29813-29825, 29816-29828, 31329-31341, 31677-31689, 31680-31692, 31732-31744, 31735-31747, 36137-36149, 36140-36152, 36812-36824, 36815-36827, 37413-37425, 38679-38691, 39474-39486, 39477-39489;

SOX5\_01: 27397-27413, 27572-27588, 28100-28116, 29230-29246, 29439-29455, 30690-30706, 31595-31611, 33871-33887, 34113-34129, 34624-34640, 37668-37684, 38582-38598, 39124-39140, 40410-40426;

SRV\_02: 20016-20032, 22410-22426, 27329-27345, 29162-29178, 29499-29515, 30646-

30662, 31503-31519, 35928-35944, 37324-37340;

TATA\_01: 32722-32738, 32729-32745, 32807-32823, 33825-33841, 34120-34136, 35433-35449, 36593-36609;

TATA\_C: 11015-11031, 11817-11833, 13635-13651, 14930-14946;

TCF11\_01: 18543-18549, 22574-22580, 31281-31297, 31489-31505, 38754-38770;

USE\_01: 23075-23087, 32577-32589;

VMYB\_02: 11526-11538, 17384-17396, 18400-18412, 19549-19561, 22188-22200, 40486-40508 and

XFD2\_01: 16620-16636, 18153-18169, 22102-22118, 23141-23157.

And a transcription binding site selected from the group consisting of

#### BINDING SITES

huMDM2, 1 location in SEQ ID NO:4

AP1\_C: 44584-44594, 49069-49079;

AP1\_Q2: 42174-42184, 45217-45227, 48422-48422, 50447-50457.

APL\_Q4: 42702-42712, 50806-50816;

AP4\_Q6: 42117-42133, 42118-42134, 42244-42260, 45432-45448; 45433-45449, 46609-46625;

BRN2\_Q1: 42310-42328, 44022-44040, 47514-47532, 48900-48918, 48967-48985;

CAAT\_Q1: 44866-44880;

CDPCR3HD\_Q1: 45671-45689, 49219-49237;

CREL\_Q1: 42437-42449, 49797-49809;

FREAC7\_Q1: 47026-47042, 47292-47308, 47658-47674;

GATA1\_Q2: 43482-43494, 48926-48938, 49284-49296;

GATA1\_Q3: 47371-47383;

GATA1\_Q4: 43054-43066, 43162-43162, 43967-43979, 45464-45476, 45916-45928, 47763-47775;

GATA1\_Q5: 49319-49331, 49459-49471;

GATA1\_Q6: 47590-47602;

GATA2_02:	42660-42672, 43475-43487;
GATA2_03:	43714-43726, 50948-50960;
GATA3_02:	49155-49167, 49844-49856;
GATA3_03:	42202-42214, 44810-44822, 48438-48450, 49136-49148, 49337-49349, 49869-49881;
GATA_C:	44011-44023, 45256-45268, 45823-45835, 47915-47927, 49201-49213, 49573-49585;
GFI1_01:	46606-46618, 47063-47075;
HFH3_01:	47030-47046, 47284-47300, 47288-47304;
IK2_01:	45275-45287;
LYF1_01:	44564-44576, 46991-47003, 49567-49579;
MAX_01:	43234-43246, 48726-48738;
MZF1_01:	41772-41780, 42290-42298, 42295-42303, 44507-44515, 45105-45113, 45203-45211, 49948-49956, 50774-50782;



NF1_Q6:	50209-50229;
NFAT_Q6:	42061-42079, 44418-44436, 46399-46417, 47974-47992, 49267-49285, 49964-49982, 50392-50410;
NKX25_01:	42394-42408, 43507-43521, 46115-46129;
RORA1_01:	45073-45091, 48718-48736;
S8_01:	43552-43564, 45214-45226, 47160-47172, 48419-48431, 49295-49307, 50379-50391;
SOX5_01:	43716-43732, 46351-46367, 47156-47172, 47774-47790, 47868-47884, 47974-47990, 48915-48931, 50323-50339;
TATA_01:	45588-45604, 47625-47641, 48026-48042, 48659-48675, 49056-49072, 49079-49095, 49152-49168;
TCF11_01:	49115-49131;
VMYB_02:	42010-42022, 42279-42291, 44651-44663; and
XFD2_01:	42870-42886, 42910-42926.

10. A composition comprising the nucleic acid molecule of claim 7 and a carrier.
15. A kit comprising the nucleic acid molecule of claim 7.
16. The kit according to claim 15, in which the nucleic acid molecule is labeled with a detectable substance.
17. A solid support comprising the nucleic acid molecule of claim 7.
18. The solid support of claim 17 wherein said support is a microarray.
20. The solid support of claim 18, which further comprises a nucleic acid molecule encoding human mouse double minute 2 homolog, complementary sequence thereof or a portion of said nucleic acid molecule containing at least 20 contiguous nucleotides.
22. A method of identifying variants of SEQ ID NO:4, or its complementary sequence, comprising  
isolating genomic DNA from a subject and determining the presence or absence of a variant in said genomic DNA using the nucleic acid molecule of claim 7.
23. A method for detecting the presence or absence of SEQ ID NO:4 or its complementary sequence in a sample, said method comprising (a) contacting the sample with the nucleic acid molecule of claim 7 and (b) determining whether the nucleic acid molecule binds to said

nucleic acid sequence in the sample.

24. An isolated nucleic acid molecule 20-5000 contiguous nucleotides in length consisting of a reverse or forward strand of a contiguous exon-intron region between nucleotides 41738-9502 of SEQ ID NO:4, or contiguous intron-exon region between nucleotides 41738-9502 of SEQ ID NO:4, wherein a sequence segment comprising 41738-9502 of SEQ ID NO:4 encodes human mouse double minute 2 homolog depicted in SEQ ID NO:2.

25. The isolated nucleic acid molecule of claim 24, wherein said nucleic acid molecule is 20-5000 contiguous nucleotides in length and comprises nucleotides 41739-41738, 40645-40646, 36309-36310, 36384-36385, 32994-32995, 33126-33127, 29564-29565, 29615-29616, 25507-25508, 25287-25288, 25383-25384, 25576-25577, 21006-21007, 21168-21169, 13953-13954, 14109-14110, 13188-13189, 13266-13267, 10664-10665 and/or 9504-9503 of SEQ ID NO:4 or their reverse strands.

30. A microarray comprising a plurality of the nucleic acid molecules of claim 7.




31. The microarray of claim 30 wherein said microarray further comprises a nucleic acid molecule encoding human mouse double minute 2 homolog, complementary sequence thereof or a portion of said nucleic acid molecule containing at least 20 contiguous nucleotides.

32. A method for detecting the presence of a nucleic acid sequence of SEQ ID NO:4 or its complementary sequence in a sample, said method comprising contacting the sample with the

nucleic acid molecule of claim 7 and determining whether the nucleic acid molecule binds to said nucleic acid sequence in the sample.

## EVIDENCE APPENDIX

1. Muzny et al. (description of sequence of AC025423 made of record by Examiner)
2. Vogelstein et al., US Patent No. 5,411,860 and NM\_002392 (made of record by examiner)
3. Watson, Recombinant DNA
4. Genecard profile of human MDM2 submitted with amendment dated August 25, 2005
5. Anderson et al., 1996, Mammalian Genome 7:780-783 submitted with amendment dated August 25, 2005
6. Bureau, 1995, Genomics 28: 109-112 submitted with amendment dated August 25, 2005
7. Information Disclosure Statement submitted March 1, 2005 (and references submitted)
8. Liang et al. submitted with Appeal Brief dated March 2, 2008
9. Wong et al., 2000, Genome Research 10: 1672-1678

Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy Bio

Search  for

Limits Preview/Index History Clipboard Details

Display  Show

**1: AC025423** Reports Homo sapiens 12 B...[gi:14578057] [Links](#)

LOCUS AC025423 150579 bp DNA linear PRI 23-JAN-2003  
 DEFINITION Homo sapiens 12 BAC RP11-61102 (Roswell Park Cancer Institute Human BAC Library): complete sequence.  
 AC025423  
 VERSION AC025423.32 GI:14578057  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)

ORGANISM **Homo sapiens**  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 150579)  
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,R.C., Arc,J.R., Banks,T., Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowles,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,B., Chen,Z., Chin,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Southwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,E., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,D.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovac,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okmonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolf,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shoostari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,b., Vera,V., Villalón,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,

h

cb

hg e e e e fcg

Washington, C., Watlington, S., Williams, G., Williamson, A.,  
 Wiecek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J.,  
 Zorrilla, S., Kucherlapati, R., Weinstock, G. and Gibbs, R.

TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 150579)  
 AUTHORS Worley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-MAR-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 150579)  
 AUTHORS Worley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-JUN-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 150579)  
 AUTHORS Worley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-SEP-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE 5 (bases 1 to 150579)  
 AUTHORS Worley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-JUN-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

COMMENT On Jun 30, 2001 this sequence version replaced gi:14575757.  
 INFORMATION: <http://www.hqsc.bcm.tmc.edu/> or email  
 gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the  
 entire insert of this clone. Overlapping regions of clones are only  
 sequenced and submitted once, so the sequence for the remainder of  
 the insert may be found in the record for the adjacent clones.  
 Overlapping clones are noted at the beginning and end of the  
 Features listing.

#### ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches  
 of a local database that includes entries from dbSTS, GDB, and  
 local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and F. Green,  
 unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST  
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
 EST and cDNA sequences. Genes demonstrate at least two exons  
 flanked by consensus splice sites that maintained sequence  
 continuity across the splice junctions. Sequences that are not  
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
 standard of double strand coverage with a minimum of 2 clones and 2  
 reads with no ambiguities or 2 chemistries with a minimum of 2  
 clones and 3 reads with no ambiguities. If the sequence quality for  
 a region does not meet this standard, it will be indicated in the  
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

#### QUALSTAT-REPORT.

FEATURES	Location/Qualifiers
source	1..159579 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="12" /clone="RP11-61102"
<u>misc_feature</u>	1..1996 /note="overlaps bases 96763..98760 of clone AC133749" /function="clone overlap"
<u>repeat_region</u>	complement(84..376) /rpt_family="AluJb" 112..380
<u>STS</u>	/standard_name="163887"
<u>repeat_region</u>	complement(395..526) /rpt_family="L2" 1610..1838
<u>repeat_region</u>	/rpt_family="AluJb" complement(1857..2163) /rpt_family="AluSx" 2353..3314
<u>repeat_region</u>	/rpt_family="MER4-internal" 3319..3635
<u>repeat_region</u>	/rpt_family="AluJo" 4067..4746
<u>repeat_region</u>	/rpt_family="LTR8A" 4779..5068
<u>repeat_region</u>	/rpt_family="AluJo" 5150..5291
<u>repeat_region</u>	/rpt_family="L2" complement(6105..6362) /rpt_family="MIR" 6808..6930
<u>repeat_region</u>	/rpt_family="L2" 7845..8088
<u>repeat_region</u>	/rpt_family="AluJb" 8089..8114
<u>repeat_region</u>	/rpt_family="(CAAA)n" 8274..8307
<u>repeat_region</u>	/rpt_family="(CA)n" 8371..8654
<u>repeat_region</u>	/rpt_family="L1MB2" 9026..9335
<u>repeat_region</u>	/rpt_family="L1MC2" 9346..9439
<u>repeat_region</u>	/rpt_family="A-rich" 9443..9666
<u>repeat_region</u>	/rpt_family="L1M1" 9973..10073
<u>repeat_region</u>	/rpt_family="LTR37B" 10074..10257
<u>repeat_region</u>	/rpt_family="(TA)n"



```

repeat_region 10289..10653
                /rpt_family="(TA)n"
repeat_region 10666..10821
                /rpt_family="LTR37B"
repeat_region complement(11068..11199)
                /rpt_family="AluJo/FRAM"
repeat_region 11364..11538
                /rpt_family="L2"
repeat_region complement(11956..12173)
                /rpt_family="MER20"
repeat_region complement(12540..12607)
                /rpt_family="L1"
repeat_region 12797..13099
                /rpt_family="AluSx"
repeat_region complement(13269..13362)
                /rpt_family="MER5A"
repeat_region 13888..14200
                /rpt_family="AluY"
repeat_region 14201..14242
                /rpt_family="(CAAA)n"
repeat_region 14243..14265
                /rpt_family="AluY"
repeat_region 14269..14580
                /rpt_family="AluSq"
repeat_region complement(14680..14784)
                /rpt_family="MIR"
repeat_region complement(14816..14853)
                /rpt_family="L1PA6"
repeat_region 14851..14973
                /rpt_family="L1PA6"
repeat_region complement(15100..15219)
                /rpt_family="MER5B"
repeat_region 15328..15385
                /rpt_family="GA-rich"
repeat_region 16238..16266
                /rpt_family="GC-rich"
repeat_region 16285..16309
                /rpt_family="GC-rich"
repeat_region complement(17545..17644)
                /rpt_family="MIR"
repeat_region complement(17704..18015)
                /rpt_family="AluSg"
repeat_region 18513..18537
                /rpt_family="AT-rich"
repeat_region 19097..19398
                /rpt_family="AluSq"
repeat_region 19419..19438
                /rpt_family="(TTTA)n"
repeat_region 20108..20173
                /rpt_family="L2"
repeat_region 20324..20346
                /rpt_family="AT-rich"
repeat_region 20471..20735
                /rpt_family="AluJb"
repeat_region complement(20866..21040)
                /rpt_family="MIR"
repeat_region complement(21081..21763)
                /rpt_family="L1PA3"
repeat_region 21763..22907
                /rpt_family="L1PA3"

```



## Sequence Revision History

Published Nucleotide Protein Genome Structure PBC Taxonomy OMIM

Find (Accessions, GI numbers or Fasta style Seqids) AC025423

About Entrez

Show difference between I and II as GenBank/GenPept

Entrez

Revision history for AC025423

Search for Genes  
LocustLink provides curated  
information for human, fruit  
fly, mouse, rat, and  
zebrafish

Help/FAQ

Batch Entrez: Upload a  
file of GI or accession  
numbers to retrieve  
protein or nucleotide  
sequences

Check sequence  
revision history

How to create WWW  
links to Entrez

LinkCut

Cubby

Related resources

Sci AR

Reference sequence project

LocustLink

Clusters of orthologous groups

Protein reviews on the web

GI	Version	Update Date	Status	I	II
14578057	32	Jan 23 2003 4:22 AM	Live	<input checked="" type="radio"/>	<input type="radio"/>
14578057	32	Sep 30 2002 11:59 AM	Dead	<input type="radio"/>	<input checked="" type="radio"/>
14578057	32	Jun 30 2001 4:17 AM	Dead	<input type="radio"/>	<input type="radio"/>
14575757	31	Jun 29 2001 4:09 AM	Dead	<input type="radio"/>	<input type="radio"/>
14547368	30	Jun 26 2001 4:45 AM	Dead	<input type="radio"/>	<input type="radio"/>
14547368	30	Jun 25 2001 4:47 AM	Dead	<input type="radio"/>	<input type="radio"/>
14532332	29	Jun 24 2001 4:19 AM	Dead	<input type="radio"/>	<input type="radio"/>
14333711	28	Jun 10 2001 4:08 AM	Dead	<input type="radio"/>	<input type="radio"/>
14333711	28	Jun 9 2001 4:15 AM	Dead	<input type="radio"/>	<input type="radio"/>
14328936	27	Jun 8 2001 4:13 AM	Dead	<input type="radio"/>	<input type="radio"/>
14327716	26	Jun 7 2001 4:09 AM	Dead	<input type="radio"/>	<input type="radio"/>
14318341	25	Jun 6 2001 4:13 AM	Dead	<input type="radio"/>	<input type="radio"/>
14149585	24	May 19 2001 4:26 AM	Dead	<input type="radio"/>	<input type="radio"/>
14030796	23	May 14 2001 4:10 AM	Dead	<input type="radio"/>	<input type="radio"/>
14029854	22	May 13 2001 4:16 AM	Dead	<input type="radio"/>	<input type="radio"/>
14028873	21	May 12 2001 4:13 AM	Dead	<input type="radio"/>	<input type="radio"/>
13959114	20	May 7 2001 4:13 AM	Dead	<input type="radio"/>	<input type="radio"/>
13957550	19	May 6 2001 9:19 AM	Dead	<input type="radio"/>	<input type="radio"/>
13957550	19	May 5 2001 4:12 AM	Dead	<input type="radio"/>	<input type="radio"/>
13940523	18	May 4 2001 4:21 AM	Dead	<input type="radio"/>	<input type="radio"/>
13928603	17	May 2 2001 4:09 AM	Dead	<input type="radio"/>	<input type="radio"/>
13661843	16	Apr 21 2001 4:14 AM	Dead	<input type="radio"/>	<input type="radio"/>
13661843	16	Apr 20 2001 4:20 AM	Dead	<input type="radio"/>	<input type="radio"/>
13661843	16	Apr 19 2001 4:18 AM	Dead	<input type="radio"/>	<input type="radio"/>
13661843	16	Apr 18 2001 2:03 PM	Dead	<input type="radio"/>	<input type="radio"/>

13605934	15	Apr 14 2001 4:22 AM	Dead	c	c
13605934	15	Apr 13 2001 4:14 AM	Dead	c	c
13605934	15	Apr 12 2001 4:10 AM	Dead	c	c
12831256	14	Feb 18 2001 4:14 AM	Dead	c	c
12831256	14	Feb 17 2001 4:26 AM	Dead	c	c
12831256	14	Feb 15 2001 4:27 AM	Dead	c	c
12829983	13	Feb 14 2001 4:34 PM	Dead	c	c
12280685	12	Jan 18 2001 4:16 AM	Dead	c	c
12248203	11	Jan 17 2001 4:16 AM	Dead	c	c
1225259	10	Jan 15 2001 4:13 AM	Dead	c	c
1225259	10	Jan 14 2001 4:13 AM	Dead	c	c
12083926	9	Jan 12 2001 4:16 AM	Dead	c	c
9966558	8	Sep 6 2000 5:30 AM	Dead	c	c
9966558	8	Sep 5 2000 5:22 AM	Dead	c	c
8699775	7	Jun 26 2000 5:15 AM	Dead	c	c
8699775	7	Jun 25 2000 7:09 AM	Dead	c	c
8164471	6	Jun 15 2000 12:10 PM	Dead	c	c
8164471	6	Jun 2 2000 7:54 AM	Dead	c	c
7985973	5	May 22 2000 7:11 AM	Dead	c	c
7407868	4	Apr 5 2000 5:11 AM	Dead	c	c
7407868	4	Apr 4 2000 7:42 AM	Dead	c	c
7264143	3	Mar 19 2000 5:35 PM	Dead	c	c
7229044	2	Mar 12 2000 5:21 AM	Dead	c	c
7229044	2	Mar 11 2000 10:40 PM	Dead	c	c
7211834	1	Mar 9 2000 5:29 AM	Dead	c	c

Accession AC025423.32 was first seen at NCBI on Mar 9 2000  
5:29 AM

Disclaimer | Write to the Help Desk  
NCBI | NLM/NIH



US005411860A

**United States Patent** [19][11] **Patent Number:** **5,411,860****Vogelstein et al.**[45] **Date of Patent:** **May 2, 1995**[54] **AMPLIFICATION OF HUMAN MDM2 GENE IN HUMAN TUMORS**

[75] Inventors: Bert Vogelstein; Kenneth W. Kinzler, both of Baltimore, Md.

[73] Assignee: The Johns Hopkins University, Baltimore, Md.

[21] Appl. No.: 903,103

[22] Filed: Jun. 23, 1992

**Related U.S. Application Data**

[63] Continuation-in-part of Ser. No. 867,840, Apr. 7, 1992, abandoned.

[51] Int. Cl.<sup>5</sup> ..... C12Q 1/68; C12P 19/34

[52] U.S. Cl. .... 438/6; 435/91.2; 436/813

[58] Field of Search ..... 435/6, 91.2; 935/77; 436/813

[56] **References Cited****PUBLICATIONS**

- Romkes et al. *Biochemistry* 30:3247-3255 (1991) "Cloning & Expression of CDNA for . . .".
- Fakhrazadeh, et al., "Tumorigenic Potential Associated with Enhanced Expression of a Gene That is Amplified in a Mouse Tumor Cell Line", *The EMBO Journal*, 10(6):1565-1569 (1991).
- Hinds, et al., "Mutant p53 DNA Clones From Human Colon Carcinomas Cooperate With Ras in Transforming Primary Rat Cells: A Comparison of the Hot Spot

Mutant Phenotypes", *Cell Growth & Differentiation*, 1:561-580 (1990).

Oliner, et al., "Amplification of a Gene Encoding a p53-Associated Protein in Human Sarcomas." *Nature*, 358:80-83 (1992).

Oliner, et al., "Oncoprotein MDM2 Conceals the Activation Domain of Tumour Suppressor p53", *Nature*, 362(6423):857-860 (1993), abstract.

Leach et al., "p53 Mutation and MDM2 Amplification in Human Soft Tissue Sarcomas," *Cancer Research*, 53:2231-2234 (1993).

Momand, et al., "The mdm-2 Oncogene Product Forms a Complex With the p53 Protein and Inhibits p53-Mediated Transactivation," *Cell*, 69:1237-1245 (1992).

Primary Examiner—Margaret Parr

Assistant Examiner—Eggerton Campbell

Attorney, Agent, or Firm—Banner, Birch, McKie & Beckett

[57] **ABSTRACT**

A human gene has been discovered which is genetically altered in human tumor cells. The genetic alteration is gene amplification and leads to a corresponding increase in gene products. Detecting that the gene, designated hMDM2, has become amplified or detecting increased expression of gene products is diagnostic of tumorigenesis. Human MDM2 protein binds to human p53 and appears to allow the cell to escape from p53-regulated growth.

[57] **ABSTRACT**

A human gene has been discovered which is genetically altered in human tumor cells. The genetic alteration is gene amplification and leads to a corresponding increase in gene products. Detecting that the gene, designated hMDM2, has become amplified or detecting increased expression of gene products is diagnostic of tumorigenesis. Human MDM2 protein binds to human p53 and appears to allow the cell to escape from p53-regulated growth.

**12 Claims, 12 Drawing Sheets**

*FIG. 1A(1)*

1 GCACCGCGCGAGCTTGGCTGCTTCTGCGGGC  
\* AG  
84 GGCCGCGACCCCTCTGACCGAGATCCTGCTG  
CGT GC GG CTCGCGCTCCCCG GAAG  
168 GTGCCCTGGCCCCGAGAGTGGAAATGATCCCC  
ACC GACACCCCTGGGGGACC TCG AT  
252 GGAGTCTTGAGGGACCCCGACTCCAAGCGC  
1  
T C G C G  
336 CCTACTGATGGTGCTGTAACCACCTCACAGA  
9 P T D G A V T T S Q  
S E A S  
G C A G C  
420 TTATTAAAGTCTGTTGGTGACAAAAAGACA  
37 L L K S V G A Q K D  
N  
A G C G G C  
504 CGATTATATGATGAGAAGCAACAACATATTG  
65 R L Y D E K Q Q H I  
G G A  
588 GTGAAAGAGCACAGGAAAATATATACCATGA  
93 V K E H R K I Y T M  
A  
GC G AC G C  
672 TCTGTGAGTGAGAACAGGTGTCACCTTGAAG  
121 S V S E N R C H L E  
L S R Q P

## FIG. 1A(2)

CTGTGTGGCCCTGTGTGTCGGAAAGATGGAGCAAGA

AGCCGC GC TTCTC TCG TCGAGCT TG ACGAC  
CTTTCGCAGCCAGGAGCACCGTCCCTCCCCGGATTA

GTCCGAA ATGCGC G AAGTAG CC T CT  
GAGGCCCAGGGCGTCGTGCTTCCGCAGTAGTCAGTC

ACCGCG TTCTCCT C GCCTC C  
GAAAACCCCGGATGGTGAGGAGCAGGCAAATGTGCA  
M C

T  
TTCCAGCTTCGGAACAAGAGACCCTGGTTAGACCAA  
I P A S E Q E T L V R P

C A A A A  
CTTATACTATGAAAGAGGTTCTTTTTTATCTTGGCC  
T Y T M K E V L F Y L G  
I I I

G C G  
TATATTGTTCAAATGATCTTCTAGGAGATTGTTTG  
V Y C S N D L L G D L F  
V

A T A G CT A G A----  
TCTACAGGAACCTTGGTAGTAGTCAATCAGCAGGAAT  
I Y R N L V V N Q Q E  
A S -

TG T C T G C CA  
GTGGGAGTGATCAAAAGGACCTTGTACAAGAGCTTC  
G G S D Q K D L V Q E L  
L P L A P

*FIG. 1A(3)*

AGCCGAGCCCGAGGGGC	83	Human nt
CATG CGCTCA G C		Mouse nt
GTGCGTACGAGCGCCA	167	Human nt
GGGCGAGC GAGACC		Mouse nt
CCCGTGAAGGAACTGG	251	Human nt
G		Mouse nt
ATACCAACATGTCTGTA	335	Human nt
N T N M S V	8	Human a.a.
		Mouse a.a.
A		Mouse nt
AGCCATTGCTTTTGAAG	419	Human nt
K P L L L K	36	Human a.a.
		Mouse a.a.
G		Mouse nt
AGTATATTATGACTAAA	503	Human nt
Q Y I M T K	64	Human a.a.
		Mouse a.a.
A C G T		Mouse nt
GCGTGCCAAGCTTCTCT	587	Human nt
G V P S F S	92	Human a.a.
		Mouse a.a.
----- T C		Mouse nt
CATCGGACTCAGGTACA	671	Human nt
S S D S G T	120	Human a.a.
- -		Mouse a.a.
CA		Mouse nt
AGGAAGAGAAACCTTCA	755	Human nt
Q E E K P S	148	Human a.a.
P		Mouse a.a.

*FIG. 1B(1)*

		TG		AA			TG	
756		TCTTCACATTTGGTTTCTAGACCATCT						
149		S S H L V S R P S						
			D		I		L	
		G G G CC G G				G GG		
840		GGTGAACGACAAAGAAAACGCCACAAA						
177		G E R Q R K R H K						
				H		R R		
		G CAGCGGCGGCACGAGCA CAGT						
924		ATATGT-----TGTGAA						
205		I C - - - - C E						
		M S G G T S S S						
		G T CC						
993		GTAAGTGAACATTAGGTGATTGGTTG						
228		V S E H S G D W L						
					C			
		G C G C						
1077		TCAGAAGATTATAGCCTTAGTGAAGAA						
256		S E D Y S L S E E						
					D			
		A A C C T						
1161		GGGGAGAGTGATACAGATTCATTGAA						
284		G E S D T D S F E						
		T C A						
1245		AATCCCCCCTTCCATCACATTGCAAC						
312		N P P L P S H C N						
					K			
		A						
1329		GAAATCTCTGAGAAAGCCAAACTGGAA						
340		E I S E K A K L E						



*FIG. 1B(2)*

T C G  
ACCTCATCTAGAAGGAGAGCAATTAGTGAGACAGAAGAA  
T S S R R R A I S E T E E  
S

----- G CCG G  
TCTGATAGTATTTCCCTTTTCCTTTGATGAAAGCCTGGCT  
S D S I S L S F D E S L A  
- - - - P G

C C C G C A C C  
AGAAGCAGTAGCAGTGAATCTACAGGGACGCCATCGAAT  
R S S S S E S T G T P S N  
S E H

T C G  
GATCAGGATTCAGTTTCAGATCAGTTTAGTGAGAATTT  
D Q D S V S D Q F S V E F

G C G G C GG  
GGACAAGAACTCTCAGATGAAGATGATGAGGTATATCAA  
G Q E L S D E D D E V Y Q  
H R

G G G T  
GAAGATCCTGAAATTTCTTAGCTGACTATTGGAAATGC  
E D P E I S L A D Y W K C  
G

C A C A C  
AGATGTTGGGCCCTTCGTGAGAATTGGCTTCCTGAAGAT  
R C W A L R E N W L P E D  
T D

G T G A A G G  
AACTCAACACAAAGCTGAAGAGGGCTTTGATGTTTCCTGAT  
N S T Q A E E G F D V P D  
A L

*FIG. 1B(3)*

CA GC C		Mouse nt
AATTCAGATGAATTATCT	839	Human nt
N S D E L S	176	Human a.a.
T P		Mouse a.a.
AGC G		Mouse nt
CTGTGTGTAATAAGGGAG	923	Human nt
L C V I R E	204	Human a.a.
E L		Mouse a.a.
A C A C		Mouse nt
CCGGATCTTGATGCTGGT	992	Human nt
P D L D A G	227	Human a.a.
Q D		Mouse a.a.
G G		Mouse nt
GAAGTTGAATCTCTCGAC	1076	Human nt
E V E S L D	255	Human a.a.
		Mouse a.a.
C A C A		Mouse nt
GTTACTGTGTATCAGGCA	1160	Human nt
V T V Y Q A	283	Human a.a.
T		Mouse a.a.
C		Mouse nt
ACTTCATGCAATGAAATG	1244	Human nt
T S C N E M	311	Human a.a.
		Mouse a.a.
G T		Mouse nt
AAAGGGAAAGATAAAGGG	1228	Human nt
K G K D K G	339	Human a.a.
V		Mouse a.a.
G C GCTG C A		Mouse nt
TGTAAAAAACTATAGTG	1412	Human nt
C K K T I V	367	Human a.a.
G L T E		Mouse a.a.

## FIG. 1C(1)

		G	T	A		C		C		G																				
1413	A	A	T	G	A	T	C	C	A	G	A	G	A	G	T	C	A	T	G	T	G	A	G	G	A					
368	N	D	S	R	E	S	C	V	E	E																				
				A	K		P		A																					
				C	A		G		C	C		G																		
1494	T	C	T	C	A	G	C	C	A	T	C	T	T	C	T	A	G	T	A	G	C	A	T	T	A	T				
395	S	Q	P	S	T	S	S	S	S	I	I																			
												V																		
				C					C	C	T	G																		
1578	G	A	A	G	A	G	A	G	T	G	T	G	G	A	A	T	C	T	A	G	T	T	G	C	C	C	T			
423	E	E	S	V	E	S	S	L	P	L																				
	D							F	S																					
				T	C		G	T		C	C	T	A																	
1662	G	T	C	A	T	G	G	C	A	A	A	C	A	G	A	C	A	T	C	T	T	A	T	G	G	C				
451	V	H	G	K	T	G	H	L	M	A																				
												S																		
				G			C																							
1746	A	G	A	C	A	C	C	A	A	T	T	C	A	A	T	G	A	T	G	T	G	C	T	A	A	C				
479	R	Q	P	I	Q	M	I	V	L	T																				
											S																			
1830	T	A	A	C	C	T	A	G	G	A	A	T	T	T	A	G	A	C	A	C	C	T	G	A	A	A	T			
1914	T	T	A	G	T	A	T	A	A	T	T	G	A	C	C	T	A	C	T	T	T	G	G	T	A	G	T	G	G	A
1998	A	C	T	C	T	A	A	T	T	T	T	T	A	A	T	A	T	T	T	C	T	A	C	T	C	T	G	T		
2082	A	T	G	T	A	A	C	T	T	A	T	T	T	T	T	T	T	T	T	T	T	G	A	G	A	C	C	G	A	G
2166	C	T	C	T	G	C	C	C	T	C	C	C	G	G	G	T	C	G	C	A	C	A	T	T	C	T	C			
2250	T	A	A	T	T	T	T	T	T	T	T	T	T	T	T	T	A	G	T	A	G	A	C	A	G	A	C	A	G	G
2334	C	T	C	G	G	C	T	C	C	C	A	A	G	T	G	C	T	G	G	G	A	T	T	A	C	A	G			

*FIG. 1C(2)*

G CAGC G G GGCCGA GA GC C TG C  
AAT---GATGATAAAATTACACAAGCTTCACAATCAC  
N - D D K I T Q A S Q S  
D S E E A E T P L

AGC G--- A  
TATAGCAGCCCAAGAAGATGTGAAAGAGTTTGAAAGGG  
Y S S Q E D V K E F E R  
S L - K

C A C C G G G  
AATGCCATTGAACCTTGTGTGATTTGTCAAGGTCGAC  
N A I E P C V I C Q G R

T C G A A C  
TGCTTTACATGTGCAAAGAAGCTAAAGAAAAGGAATA  
C F T C A K K L K K R N

C AA C CTCA A A T  
TATTTCCCCTAGTTGACCTG---TCTATAAGAGAATT  
Y F P  
N

TATTCACATATATCAAAGTGAGAAAATGCCTCAATTC  
ATAGTGAATACTTACTATAATTTGACTTGAATATGTA  
CTTAAATGAGAAGTACTTGGTTTTTTTTTTCTTAAAT  
TCTTGCTCTGTTACCCAGGCTGGAGTGCAGTGGGTGA  
CTGCCTCAGCCTCCCAATTAGCTTGGCCTACAGTCAT  
TTTCACCGTGTAGCCAGGATGGTCTCGATCTCCTGA  
CATGAGCCACCG

## FIG. 1C(3)

G G C		Mouse nt
AAGAAAGTGAAGACTAT	1493	Human nt
Q E S E D Y	394	Human a.a.
D		Mouse a.a.
G G GC		Mouse nt
AAGAAACCCAAGACAAA	1577	Human nt
E E T Q D K	422	Human a.a.
H		Mouse a.a.
C		Mouse nt
CTAAAAATGGTTGCATT	1661	Human nt
P K N G C I	450	Human a.a.
		Mouse a.a.
G C		Mouse nt
AGCCCTGCCCAGTATGT	1745	Human nt
K P C P V C	478	Human a.a.
		Mouse a.a.
T *		Mouse nt
ATATATTTCTAACTATA	1829	Human nt
	491	Human a.a.
		Mouse a.a.
ACATAGATTTCTTCTCT	1913	Human nt
GCTCATCCTTTACACCA	1997	Human nt
ATGTATATGACATTTAA	2081	Human nt
TCTTGGCTCACTGCAAG	2165	Human nt
CTGCCACCACACCTGGC	2249	Human nt
CCTCGTGATCCGCCAC	2333	Human nt
	2372	Human nt

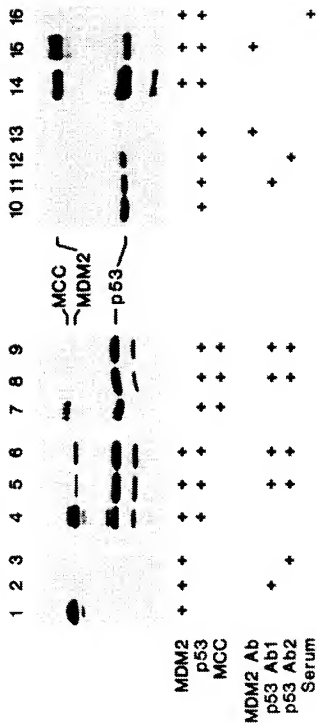
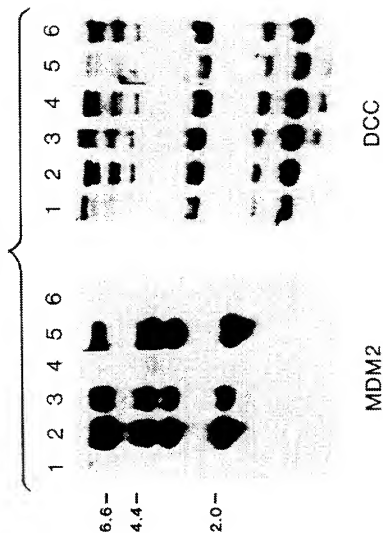
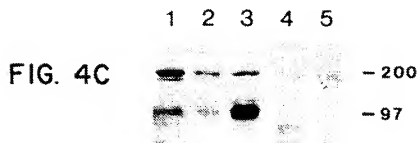
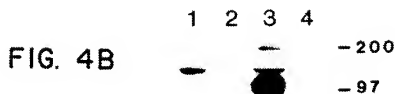
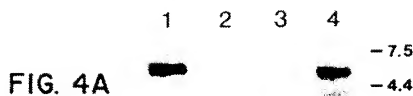


FIG. 3







# AMPLIFICATION OF HUMAN MDM2 GENE IN HUMAN TUMORS

This application is a continuation-in-part of U.S. application Ser. No. 07/667,840, filed on Apr. 7, 1992, which is now abandoned.

## FIELD OF THE INVENTION

The invention relates to the area of cancer diagnosis and therapeutics. More particularly, the invention relates to the detection of a gene which is amplified in certain human tumors.

## BACKGROUND OF THE INVENTION

According to the Knudson model for tumorigenesis (Cancer Research, 1985, vol. 45, p. 1482), there are tumor suppressor genes in all normal cells which, when they become non-functional due to mutation, cause neoplastic development. Evidence for this model has been found in cases of retinoblastoma and colorectal tumors. The implicated suppressor genes in these tumors, RB and p53 respectively, were found to be deleted or altered in many of the tumors studied.

The p53 gene product, therefore, appears to be a member of a group of proteins which regulate normal cellular proliferation and suppression of cellular transformation. Mutations in the p53 gene have been linked to tumorigenesis, suggesting that alterations in p53 protein function are involved in cellular transformation. The inactivation of the p53 gene has been implicated in the genesis or progression of a wide variety of carcinomas (Nigro et al., 1989, Nature 342:705-708), including human colorectal carcinomas (Baker et al., 1989, Science 244:217-221), human lung cancer (Takahashi et al., 1989, Science 246:491-494; Iggo et al., 1990, Lancet 335:675-679), chronic myelogenous leukemia (Kelman et al., 1989, Proc. Natl. Acad. Sci. USA 86:6783-6787) and osteogenic sarcomas (Masuda et al., 1987, Proc. Natl. Acad. Sci. USA 84:7716-7719).

While there exists an enormous body of evidence linking p53 gene mutations to human tumorigenesis (Hollstein et al., 1991, Science 253:49-53) little is known about cellular regulators and mediators of p53 function.

Hinds et al. (Cell Growth & Differentiation, 1990, 1:571-580), found that p53 cDNA clones, containing a point mutation at amino acid residue 143, 175, 273 or 281, cooperated with the activated ras oncogene to transform primary rat embryo fibroblasts in culture. These mutant p53 genes are representative of the majority of mutations found in human cancer. Hollstein et al., 1991, Science 253:49-53. The transformed fibroblasts were found to produce elevated levels of human p53 protein having extended half-lives (1.5 to 7 hours) as compared to the normal (wild-type) p53 protein (20 to 30 minutes).

Mutant p53 proteins with mutations at residue 143 or 175 form an oligomeric protein complex with the cellular heat shock protein hsc70. While residue 273 or 281 mutants do not detectably bind hsc70, and are poorer at producing transformed foci than the 175 mutant, complex formation between mutant p53 and hsc70 is not required for p53-mediated transformation. Complex formation does, however, appear to facilitate this function. All cell lines transformed with the mutant p53 genes are tumorigenic in athymic (nude) mice. In contrast, the wild-type human p53 gene does not possess

transforming activity in cooperation with ras. Tuck and Crawford, 1989, Oncogene Res. 4:81-96.

Hinds et al. supra also expressed human p53 protein in transformed rat cells. When the expressed human p53 was immunoprecipitated with two p53 specific antibodies directed against distinct epitopes of p53, an unidentified Mr 90,000 protein was coimmunoprecipitated. This suggested that the rat Mr 90,000 protein is in a complex with the human p53 protein in the transformed rat cell line.

As mentioned above, levels of p53 protein are often higher in transformed cells than normal cells. This is due to mutations which increase its metabolic stability (Owen et al., 1981, Mol. Cell. Biol. 1:101-110; Reich et al., (1983), Mol. Cell. Biol. 3:2143-2150). The stabilization of p53 has been associated with complex formation between p53 and viral or cellular proteins. (Linzer and Levine, 1979, Cell 17:43-52; Crawford et al., 1981, Proc. Natl. Acad. Sci. USA 78:41-45; Dippold et al., 1981, Proc. Natl. Acad. Sci. USA 78:1695-1699; Lane and Crawford, 1979, Nature (London) 278:261-263; Hinds et al., 1987, Mol. Cell. Biol. 7:2863-2869; Finlay et al., 1988, Mol. Cell. Biol. 8:531-539; Sarnow et al., 1982, Cell 28:387-394; Gronostajski et al., 1984, Mol. Cell. Biol. 4:442-448; Pinhasi-Kimhi et al., 1986, Nature (London) 320:182-185; Rusconi and Scolnick, 1983, J. Virol. 46:1022-1026; Pinhasi and Oren, 1984, Mol. Cell. Biol. 4:2180-2186; and Sturzbecher et al., 1987, Oncogene 1:201-211.) For example, p53 protein has been observed to form oligomeric protein complexes with the SV40 large T antigen, the adenovirus type 5 E1B-Mp 55,000 protein, and the human papilloma virus type 16 or 18 E6 product. Linzer and Levine, 1979, Cell 17:43-52; Lane and Crawford, 1979, Nature, 278:261-263; Sarnow et al., 1982, Cell 28:387-394; Werness et al., 1990, Science, 248:76-79. Similarly, complexes have been observed of p105<sup>ras</sup> (the product of the retinoblastoma susceptibility gene) with T antigen (DeCaprio et al., 1988, Cell 54:275-283), the adenovirus E1A protein (Whyte et al., 1988, Nature 334:124-129) and the E7 protein of human papilloma virus 16 or 18 (Münzer et al., 1989, EMBO J. 8:4099-4105). It has been suggested that interactions between these viral proteins and p105<sup>ras</sup> inactivate a growth-suppressive function of p105<sup>ras</sup>, mimicking deletions and mutations commonly found in the RB gene in tumor cells. In a similar fashion, oligomeric protein complex formation between these viral proteins and p53 may eliminate or alter the function of p53. Finlay et al., 1989, Cell 57:1083-1093.

Fakhrazadeh et al. (EMBO J. 1991, 10:1565-1569) analyzed amplified DNA sequences present in a tumorigenic mouse cell line (i.e., 3T3DM, a spontaneously transformed derivative of mouse Balb/c cells). Studies were conducted to determine whether any of the amplified genes induced tumorigenicity following introduction of the amplified genes into a nontransformed recipient cell (e.g., mouse NIH3T3 or Rat2 cells). The resulting cell lines were tested for tumorigenicity in nude mice. A gene, designated MDM2, which is amplified more than 50-fold in 3T3DM cells, induced tumorigenicity when overexpressed in NIH3T3 and Rat 2 cells. From the nucleotide and predicted amino acid sequence of mouse MDM2 (mMDM2), Fakhrazadeh speculated that this gene encodes a potential DNA binding protein that functions in the modulation of expression of other genes and, when present in excess, interferes with normal constraints on cell growth.

## SUMMARY OF THE INVENTION

It is an object of the invention to provide a method for diagnosing a neoplastic tissue, such as sarcoma, in a human.

It is another object of the invention to provide a cDNA molecule encoding the sequence of human MDM2.

Yet another object of the invention is to provide a preparation of human MDM2 protein which is substantially free of other human cellular proteins.

Still another object of the invention is to provide DNA probes capable of hybridizing with human MDM2 genes or mRNA molecules.

Another object of the invention is to provide antibodies immunoreactive with human MDM2 protein.

Still another object of the invention is to provide kits for detecting amplification or elevated expression of human MDM2.

Yet another object of the invention is to provide methods for identifying compounds which interfere with the binding of human MDM2 to human p53.

A further object of the invention is to provide a method of treating a neoplastic human cell.

It has now been discovered that hMDM2, a heretofore unknown human gene, plays a role in human cancer. The hMDM2 gene has been cloned and the recombinant derived hMDM2 protein shown to bind to human p53 *in vitro*. hMDM2 has been found to be amplified in some neoplastic cells and the expression of hMDM2-encoded products has been found to be correspondingly elevated in tumors with amplification of this gene. The elevated levels of MDM2 appear to sequester p53 and allow the cell to escape from p53-regulated growth.

## BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A-C shows the cDNA sequence of human MDM2. In this figure, human and mouse nucleotide and amino acid sequences are compared, the mouse sequence being shown only where it differs from the corresponding human sequence.

FIG. 2 shows that hMDM2 binds to p53.

FIG. 3 illustrates the amplification of the hMDM2 gene in sarcomas.

FIG. 4A-C illustrates hMDM2 expression.

## DETAILED DESCRIPTION OF THE INVENTION

It is a discovery of the present invention that a gene exists which is amplified in some human tumors. The amplification of this gene, designated MDM2, is diagnostic of neoplasia or the potential therefor. Detecting the elevated expression of human MDM2-encoded products is also diagnostic of neoplasia or the potential for neoplastic transformation. Over a third of the sarcomas surveyed, including the most common bone and soft tissue forms, were found to have amplified hMDM2 sequences. Expression of hMDM2 was found to be correspondingly elevated in tumors with the gene amplification.

Other genetic alterations leading to elevated hMDM2 expression may be involved in tumorigenesis also, such as mutations in regulatory regions of the gene. Elevated expression of hMDM2 may also be involved in tumors other than sarcomas including but not limited to those in which p53 inactivation has been implicated. These

include c., rectal carcinoma, lung cancer and chronic myelogenous leukemia.

According to one embodiment of the invention, a method of diagnosing a neoplastic tissue in a human is provided. Tissue or body fluid is isolated from a human, and the copy number of human MDM2 genes is determined. Alternatively, expression levels of human MDM2 gene products can be determined. These include protein and mRNA.

Body fluids which may be tested include urine, serum, blood, feces, saliva, and the like. Tissues suspected of being neoplastic are desirably separated from normal appearing tissue for analysis. This can be done by paraffin or cry, star sectioning or flow cytometry, as is known in the art. Failure to separate neoplastic from non-neoplastic cells can confound the analysis. Adjacent non-neoplastic tissue or any normal tissue can be used to determine a base-line level of expression or copy number, against which the amount of hMDM2 gene or gene products can be compared.

The human MDM2 gene is considered to be amplified if the cell contains more than the normal copy number (2) of this gene per genome. The various techniques for detecting gene amplification are well known in the art. Gene amplification can be determined, for example, by Southern blot analysis, as described in Example 4, wherein cellular DNA from a human tissue is digested, separated, and transferred to a filter where it is hybridized with a probe containing complementary nucleic acids. Alternatively, quantitative polymerase chain reaction (PCR) employing primers can be used to determine gene amplification. Appropriate primers will bind to sequences that bracket human MDM2 coding sequences. Other techniques for determining gene copy number as are known in the art can be used without limitation.

The gene product which is measured may be either mRNA or protein. The term elevated expression means an increase in mRNA production or protein production over that which is normally produced by non-cancerous cells. Although amplification has been observed in human sarcomas, other genetic alterations leading to elevated expression of MDM2 may be present in these or other tumors. Other tumors include those of lung, breast, brain, colorectal, bladder, prostate, liver, skin, and stomach. These, too, are contemplated by the present invention. Non-cancerous cells for use in determining base-line expression levels can be obtained from cells surrounding a tumor, from other humans or from human cell lines. Any increase can have diagnostic value, but generally the mRNA or protein expression will be elevated at least about 3-fold, 5-fold, and in some cases up to about 100-fold over that found in non-cancerous cells. The particular technique employed for detecting mRNA or protein is not critical to the practice of the invention. Increased production of mRNA or protein may be detected, for example, using the techniques of Northern blot analysis or Western blot analysis, respectively, as described in Example 4 or other known techniques such as ELISA, immunoprecipitation, RIA and the like. These techniques are also well known to the skilled artisan.

According to another embodiment of the invention, nucleic acid probes or primers for the determining of human MDM2 gene amplification or elevated expression of mRNA are provided. The probe may comprise ribo- or deoxyribonucleic acids and may contain the entire human MDM2 coding sequence, a sequence com-

plementary thereto, or fragments thereof. A probe may contain, for example, nucleotides 1-949, or 1-2372 as shown in FIG. 1. Generally, probes or primers will contain at least about 14 contiguous nucleotides of the human sequence but may desirably contain about 40, 50 or 100 nucleotides. Probes are typically labelled with a fluorescent tag, a radioisotope, or the like to render them easily detectable. Preferably the probes will hybridize under stringent hybridization conditions. Under such conditions they will not hybridize to mouse MDM2. The probes of the invention are complementary to the human MDM2 gene. This means that they share 100% identity with the human sequence.

hMDM2 protein can be produced, according to the invention, substantially free of other human proteins. Provided with the DNA sequence, those of skill in the art can express the cDNA in a non-human cell. Lysates of such cells provide proteins substantially free of other human proteins. The lysates can be further purified, for example, by immunoprecipitation, coprecipitation with p53, or by affinity chromatography.

The antibodies of the invention are specifically reactive with hMDM2 protein. Preferably, they do not cross-react with MDM2 from other species. They can be polyclonal or monoclonal, and can be raised against native hMDM2 or a hMDM2 fusion protein or synthetic peptide. The antibodies are specifically immunoreactive with hMDM2 epitopes which are not present on other human proteins. Some antibodies are reactive with epitopes unique to human MDM2 and not present on the mouse homolog. The antibodies are useful in conventional analyses, such as Western blot analysis, ELISA and other immunological assays for the detection of proteins. Techniques for raising and purifying polyclonal antibodies are well known in the art, as are techniques for preparing monoclonal antibodies. Antibody binding can be determined by methods known in the art, such as use of an enzyme-labelled secondary antibody, staphylococcal protein A, and the like.

According to another embodiment of the invention, interference with the expression of MDM2 provides a therapeutic modality. The method can be applied in vivo, in vitro, or ex vivo. For example, expression may be downregulated by administering triple-strand forming or antisense oligonucleotides which bind to the hMDM2 gene or mRNA, respectively, and prevent transcription or translation. The oligonucleotides may interact with unprocessed pre-mRNA or processed mRNA. Small molecules and peptides which specifically inhibit MDM2 expression can also be used. Similarly, such molecules which inhibit the binding of MDM2 to p53 would be therapeutic by alleviating the sequestration of p53.

Such inhibitory molecules can be identified by screening for interference of the hMDM2/p53 interaction where one of the binding partners is bound to a solid support and the other partner is labeled. Antibodies specific for epitopes on hMDM2 or p53 which are involved in the binding interaction will interfere with such binding. Solid supports which may be used include any polymers which are known to bind proteins. The support may be in the form of a filter, column packing matrix, beads, and the like. Labeling of proteins can be accomplished according to any technique known in the art. Radiolabels, enzymatic labels, and fluorescent labels can be used advantageously. Alternatively, both hMDM2 and p53 may be in solution and bound molecules separated from unbound subsequently. Any separation

technique known in the art may be employed, including immunoprecipitation or immunoaffinity separation with an antibody specific for the unlabeled binding partner.

A cDNA molecule containing the coding sequence of hMDM2 can be used to produce probes and primers. In addition, it can be expressed in cultured cells, such as *E. coli*, to yield preparations of hMDM2 protein substantially free of other human proteins. The proteins produced can be purified, for example, with immunofluorescence techniques using the antibodies described above.

Kits are provided which contain the necessary reagents for determining gene copy number, such as probes or primers specific for the hMDM2 gene, as well as written instructions. The instructions can provide calibration curves to compare with the determined values. Kits are also provided to determine elevated expression of mRNA (i.e., containing probes) or hMDM2 protein (i.e., containing antibodies). Instructions will allow the tester to determine whether the expression levels are elevated. Reaction vessels and auxiliary reagents such as chromogens, buffers, enzymes, etc. may also be included in the kits.

The human MDM2 gene has now been identified and cloned. Recombinant derived hMDM2 has been shown to bind to human p53. Moreover, it has been found that hMDM2 is amplified in some sarcomas. The amplification leads to a corresponding increase in MDM2 gene products. Such amplification is associated with the process of tumorigenesis. This discovery allows specific assays to be performed to assess the neoplastic or potential neoplastic status of a particular tissue.

The following examples are provided to exemplify various aspects of the invention and are not intended to limit the scope of the invention.

#### EXAMPLE 1

To obtain human cDNA clones, a cDNA library was screened with a murine MDM2 (mMDM2) cDNA probe. A cDNA library was prepared by using polyadenylated RNA isolated from the human colonic carcinoma cell line CaCo-2 as a template for the production of random hexamer primed double stranded cDNA. Gubler and Hoffmann, 1983, *Gene* 25:263-268. The cDNA was ligated to adapters and then to the lambda YES phage vector, packaged, and plated as described by Elledge et al. (*Proc. Natl. Acad. Sci. USA*, 1991, 88:1731-1735). The library was screened initially with a <sup>32</sup>P-labelled (Feinberg and Vogelstein, 1983, *Anal. Biochem.* 132:6-13) mMDM2 cDNA probe (nucleotides 259 to 1508 (Fakhrazadeh et al., 1991, *EMBO J.* 10:1565-1569)) and then rescreened with an hMDM2 cDNA clone containing nucleotides 40 to 702.

Twelve clones were obtained, and one of the clones was used to obtain thirteen additional clones by rescreening the same library. In total, twenty-five clones were obtained, partially or totally sequenced, and mapped. Sequence analysis of the twenty-five clones revealed several cDNA forms indicative of alternative splicing. The sequence shown in FIG. 1 is representative of the most abundant class and was assembled from three clones: c14-2 (nucleotides 1-949), c89 (nucleotides 467-1737), and c33 (nucleotides 390-2372). The 3' end of the untranslated region has not yet been cloned in mouse or human. The 3' end is likely to be at or near nucleotide 1. There was an open reading frame extending from the 5' end of the human cDNA sequence to nucleotide 1784. Although the signal for translation

initiation could not be unambiguously defined, the ATG at nucleotide 312 was considered the most likely position for several reasons. First, the sequence similarity between hMDM2 and mMDM2 fell off dramatically upstream of nucleotide 312. This lack of conservation in an otherwise highly conserved protein suggested that the sequences upstream of the divergence may not code for protein. Second, an anchored polymerase chain reaction (PCR) approach was employed in an effort to acquire additional upstream cDNA sequence. Ochman et al., 1985, in PCR Technology: Principles and Applications for DNA Amplification (Erlrich, ed.), pp. 105-111 (Stockton, New York). The 5' ends of the PCR derived clones were very similar (within 3 bp) to the 5' ends of clones obtained from the cDNA library, suggesting that the 5' end of the hMDM2 sequence shown in FIG. 1 may represent the 5' end of the transcript. Third, in vitro translation of the sequence shown in FIG. 1, beginning with the methionine encoded by the nucleotide 312 ATG, generated a protein similar in size to that observed in human cells.

In FIG. 1, hMDM2 and mMDM2 nucleotide and amino acid sequences are compared. The mouse sequence is only shown where it differs from the corresponding human sequence. Asterisks mark the 5' and 3' boundaries of the previously published mMDM2 cDNA. Fakhrazadeh et al., 1991, EMBO J. 10:1565-1569. Dashes indicate insertions. The mouse and human amino acid sequences are compared from the putative translation start site at nucleotide 312 through the conserved stop codon at nucleotide 1784.

Comparison of the human and mouse MDM2 coding regions revealed significant conservation at the nucleotide (80.3%) and amino acid (80.4%) levels. Although hMDM2 and mMDM2 bore little similarity to other genes recorded in current databases, the two proteins shared several motifs. These included a basic nuclear localization signal (Tanaka, 1990, FEBS Letters 271:4146) at codons 181 to 185, several casein kinase II serine phosphorylation sites (Pinna, 1990, Biochem. et. Biophys. Acta. 1054:267-284) at codons 166 to 169, 192 to 195, 269 to 272, and 290 to 293, an acidic activation domain (Ptashne, 1988, Nature 335:683-689) at codons 223 to 274, and two metal binding sites (Harrison, 1991, Nature 353:715) at codons 305 to 322 and 461 to 478, neither of which is highly related to known DNA binding domains. The protein kinase A domain noted in mMDM2 (Fakhrazadeh et al., 1991, EMBO J. 10:1565-1569) was not conserved in hMDM2.

#### EXAMPLE 2

To determine whether the hMDM2 protein could bind to human p53 protein in vitro, an hMDM2 expression vector was constructed from the cDNA clones. The hMDM2 expression vector was constructed in pBlueScript SK+ (Stratagene) from overlapping cDNA clones. The construct contained the sequence shown in FIG. 1 from nucleotide 312 to 2176. A 42 bp black beetle virus ribosome entry sequence (Dashmashatra et al., 1987, Nucleic Acid Research 15:3933) was placed immediately upstream of this hMDM2 sequence in order to obtain a high level of expression. This construct, as well as p53 (El-Deiry et al., 1992, Nature Genetics, in press) and MCC (Kinzler et al., 1991, Science 251:1366-1370) constructs in pBlueScript SK+, were transcribed with T7 RNA polymerase and translated in a rabbit reticulocyte lysate (Promega) according to the manufacturer's instructions.

Although the predicted size of the protein generated from the construct was only 55.2 kd (extending from the methionine at nucleotide 312 to nucleotide 1784), in vitro translated protein migrated at approximately 95 kilodaltons.

Ten  $\mu$ l of lysate containing the three proteins (hMDM2, p53 and MCC), alone or mixed in pairs, were incubated at 37° C. for 15 minutes. One microgram (10  $\mu$ l) of p53 Ab1 (monoclonal antibody specific for the C-terminus of p53) or Ab2 (monoclonal antibody specific for the N-terminus of p53) (Oncogene Science), or 5  $\mu$ l of rabbit serum containing MDM2 Ab (polyclonal rabbit anti-hMDM2 antibodies) or preimmune rabbit serum (obtained from the rabbit which produced the hMDM2 Ab), were added as indicated. The polyclonal rabbit antibodies were raised against an *E. coli*-produced hMDM2-glutathione S-transferase fusion protein containing nucleotides 390 to 816 of the hMDM2 cDNA. Ninety  $\mu$ l of RIPA buffer (10 mM Tris [pH 7.5], 1% sodium deoxycholate, 1% NP40, 150 mM NaCl, 0.1% SDS), SNNTTE buffer (Levin and George, 1992, submitted for publication), or Binding Buffer (El-Deiry et al., 1992, Nature Genetics, in press) were then added and the mixtures allowed to incubate at 4° C. for 2 hours.

Two milligrams of protein A sepharose were added to each tube, and the tubes were rotated end-over-end at 4° C. for 1 hour. After pelleting and washing, the immunoprecipitates were subjected to SDS-polyacrylamide gel electrophoresis and the dried gels autoradiographed for 10 to 60 minutes in the presence of Enhance (New England Nuclear).

FIG. 2 shows the co-precipitation of hMDM2 and p53. The three buffers produced similar results, although the co-precipitation was less efficient in SNNTTE buffer containing 0.5M NaCl (FIG. 2, lanes 5 and 8) than in Binding Buffer containing 0.1M NaCl (FIG. 2, lanes 6 and 9).

In vitro translated hMDM2, p53 and MCC proteins were mixed as indicated above and incubated with p53 Ab1, p53 Ab2, hMDM2 Ab, or preimmune serum. Lanes 1, 4, 7, 10 and 14 contain aliquots of the protein mixtures used for immunoprecipitation. The bands running slightly faster than p53 are polypeptides produced from internal translation initiation sites.

The hMDM2 protein was not immunoprecipitated with monoclonal antibodies to either the C-terminal or N-terminal regions of p53 (FIG. 2, lanes 2 and 3). However, when in vitro translated human p53 was mixed with the hMDM2 translation product, the anti-p53 antibodies precipitated hMDM2 protein along with p53, demonstrating an association in vitro (FIG. 2, lanes 5 and 6). As a control, a protein of similar electrophoretic mobility from another gene (MCC (Kinzler et al., 1991, Science 251:1366-1370)) was mixed with p53. No co-precipitation of the MCC protein was observed (FIG. 2, lanes 8 and 9). When an in vitro translated mutant form of p53 (175<sup>aa</sup>) was mixed with hMDM2 protein, a similar co-precipitation of hMDM2 and p53 proteins was also observed.

In the converse of the experiments described above, the anti-hMDM2 antibodies immunoprecipitated p53 when mixed with hMDM2 protein (FIG. 2, lane 15) but failed to precipitate p53 alone (FIG. 5, lane 13). Preimmune rabbit serum failed to precipitate either hMDM2 or p53 (FIG. 2, lane 16).

## EXAMPLE 3

In order to ascertain the chromosomal localization of hMDM2, somatic cell hybrids were screened with an hMDM2 cDNA probe. A human-hamster hybrid containing only human chromosome 12 was found to hybridize to the probe. Screening of hybrids containing portions of chromosome 12 (Ture-Carel et al., 1986, *Cancer Genet. Cytogenet.* 23:291-299) with the same probe narrowed the localization to chromosome 12q12-14.

## EXAMPLE 4

Previous studies have shown that this region of chromosome 12 is often aberrant in human sarcomas. Mandahl et al., 1987, *Genes Chromosomes & Cancer* 1:9-14; Ture-Carel et al., 1986, *Cancer Genet. Cytogenet.* 23:291-299; Melitzer et al., 1991, *Cell Growth & Differentiation* 2:493-501. To evaluate the possibility that hMDM2 was genetically altered in such cancers, Southern blot analysis was performed.

FIG. 3 shows examples of the amplification of the hMDM2 gene in sarcomas. Cellular DNA (5 µg) was digested with EcoRI, separated by agarose gel electrophoresis, and transferred to nylon as described by Reed and Mann (*Nucl. Acids Res.*, 1985, 13:7207-7215). The cellular DNA was derived from five primary sarcomas (lanes 1-4, 6) and one sarcoma cell line (Osa-C1, lane 5). The filters were then hybridized with an hMDM2 cDNA fragment probe nucleotide 1-949 (see FIG. 1), or to a control probe which identifies fragments of similar size (DCC gene, 1.65 cDNA fragment). Fearon, 1989, *Science* 247:49-56. Hybridization was performed as described by Vogelstein et al. (*Cancer Research*, 1987, 47:4806-4813). A striking amplification of hMDM2 sequences was observed in several of these tumors. (See FIG. 3, lanes 2, 3 and 5). Of 47 sarcomas analyzed, 17 exhibited hMDM2 amplification ranging from 5 to 50 fold. These tumors included 7 to 13 liposarcomas, 7 of 22 malignant fibrous histiocytomas (MFH), 3 of 11 osteosarcomas, and 0 and 1 rhabdomyosarcomas. Five benign soft tissue tumors (lipomas) and twenty-seven carcinomas (colorectal or gastric) were also tested by Southern blot analysis and no amplification was observed.

## EXAMPLE 5

This example illustrates that gene amplification is associated with increased expression.

FIG. 4A illustrates hMDM2 expression as demonstrated by Northern blot analysis. Because of RNA degradation in the primary sarcomas, only the cell lines could be productively analyzed by Northern blot. RNA was separated by electrophoresis in a MOPS-formaldehyde gel and electrophoretically transferred to nylon filters. Transfer and hybridization were performed as described by Kinzler et al. (*Nature*, 1988, 332:371-374). The RNA was hybridized to the hMDM2 fragment described in FIG. 3. Ten µg of total RNA derived, respectively, from two sarcoma cell lines (Osa-CL, lane 1 and RC13, lane 2) and the colorectal cancer cell line (CaCo-2) used to make the cDNA library (lane 3). Lane 4 contains 10 µg of polyadenylated CaCo-2 RNA. RNA sizes are shown in kb. In the one available sarcoma cell line with hMDM2 amplification, a single transcript of approximately 5.5 kb was observed (FIG. 4A, lane 1). The amount of this transcript was much higher than in a sarcoma cell line without amplification

(FIG. 4A, lane 2) or in a carcinoma cell line (FIG. 4A, lane 3). When purified mRNA (rather than total RNA) from the carcinoma cell line was used for analysis, an hMDM2 transcript of 5.5 kb could also be observed (FIG. 4A, lane 4).

FIG. 4B illustrates hMDM2 expression as demonstrated by Western blot analysis of the sarcoma cell lines RC13 (lane 1), Osa-CL (lane 3), HOS (lane 4), and the carcinoma cell line CaCo-2 (lane 2).

FIG. 4C illustrates hMDM2 expression as demonstrated by Western blot analysis of primary sarcomas. Lanes 1 to 3 contain protein from sarcomas with hMDM2 amplifications, and lanes 4 and 5 contain protein from sarcomas without hMDM2 amplification.

Western blots using affinity purified MDM2 Ab were performed with 50 µg protein per lane as described by Kinzler et al. (*Mol. Cell. Biol.*, 1990, 10:634-642), except that the membranes were blocked in 10% nonfat dried milk and 10% goat serum, and secondary antibodies were coupled to horseradish peroxidase, permitting chemiluminescent detection (Amersham ECL). MDM2 Ab was affinity purified with a pATH-hMDM2 fusion protein using methods described in Kinzler et al. (*Mol. Cell. Biol.*, 1990, 10:634-642). Nonspecifically reactive proteins of 85, 120 and 200 kd were observed in all lanes, irrespective of hMDM2 amplification status. hMDM2 proteins, of 97 kd, were observed only in the hMDM2-amplified tumors. Protein marker sizes are shown in kd.

A protein of approximately 97 kilodaltons was expressed at high levels in the sarcoma cell line with hMDM2 amplification (FIG. 4B, lane 3), whereas no expression was evident in two carcinoma cell lines without amplification or in the carcinoma cell line (FIG. 4B, lanes 1, 2 and 4). Five primary sarcomas were also examined by Western blot analysis. Three primary sarcomas with amplification expressed the same size protein as that observed in the sarcoma cell line (FIG. 4C, lanes 1-3), while no protein was observed in the two sarcomas without amplification (FIG. 4C, lanes 4 and 5).

Expression of the hMDM2 RNA in the sarcoma with amplification was estimated to be at least 30 fold higher than that in the other lines examined. This was consistent with the results of Western blot analysis.

The above examples demonstrate that hMDM2 binds to p53 in vitro and is genetically altered (i.e., amplified) in a significant fraction of sarcomas, including MFH, liposarcomas, and osteosarcomas. These are the most common sarcomas of soft tissue and bone. Weiss and Enzinger, 1978, *Cancer* 41:2250-2266; Malawer et al., 1985, In: *Cancer: Principles and Practice of Oncology*, DeVita et al., Eds., pp. 1293-1342 (Lippincott, Pa.).

Human MDM2 amplification is useful for understanding the pathogenesis of these often lethal cancers.

MDM2 may functionally activate p53 in ways similar to those employed by virally encoded oncoproteins such as SV40 T-antigen, adenovirus E1B, and HPV E6. Lane and Beach, 1990, *Genes and Development* 4:1-8; Werner et al., 1990, *Science* 248:76. Consistent with this hypothesis, no sarcomas with hMDM2 amplification had any of the p53 gene mutations that occur commonly in other tumors. hMDM2 amplification provides a parallel between viral carcinogenesis and the naturally occurring genetic alterations underlying sporadic human cancer. The finding that expression of hMDM2 is correspondingly elevated in tumors with amplification of the gene is consistent with the finding that MDM2 binds to p53, and with the hypothesis that

overexpression of MDM2 in sarcomas allows escape from p53 regulated growth control. This mechanism of tumorigenesis has striking parallels to that previously observed for virally induced tumors (Lane and Be-

chinol, 1990, Genes and Development 4:1-8; Werner et al., 1990, Science 248:76), in which viral oncogene products bind to and functionally inactivate p53.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(1.1) NUMBER OF SEQUENCES: 4

## (2) INFORMATION FOR SEQ ID NO:1

(2.1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(2.2) MOLECULE TYPE: cDNA

(2.3) HYPOTHETICAL: NO

(2.4) ANTI-SENSE: NO

(2.5) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (B) CELL LINE: CaCo-2

(2.6) POSITION IN GENOME:

- (B) MAP POSITION: 12q12-14

(2.7) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 312-1734

(2.8) SEQUENCE DESCRIPTION: SEQ ID NO:1

```

GCACCGCGCG AGCTTGCGCTG CTCTGGGGCG CTGTGTGTGG CTAAGATGGA 60
      (A) LENGTH: 132 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: double
      (D) TOPOLOGY: linear
(2.2) MOLECULE TYPE: cDNA
(2.3) HYPOTHETICAL: NO
(2.4) ANTI-SENSE: NO
(2.5) ORIGINAL SOURCE:
      (A) ORGANISM: Homo sapiens
      (B) CELL LINE: CaCo-2
(2.6) POSITION IN GENOME:
      (B) MAP POSITION: 12q12-14
(2.7) FEATURE:
      (A) NAME/KEY: CDS
      (B) LOCATION: 312-1734
(2.8) SEQUENCE DESCRIPTION: SEQ ID NO:1

GTA ACC ACC TCA CAG ATT CCA GCT TCG GAA CAA GAG ACC CTG GTT AGA 398
Val Thr Thr Ser Gln Ile Pro Ala Ser Gln Gln Gln Thr Leu Val Arg
15 20 25
CCA AAG CCA TTG CTT TTA AAG TTA TTA AAG TCT GTT GGT GCA GAA AAA 444
Pro Lys Pro Leu Leu Leu Lys Leu Leu Lys Ser Val Gln Ala Gln Lys
30 35 40 45
GAC ACT TAT ACT AIG AAA GAG GAT CTT TTT TAT CTT GGC CAG TAT ATT 494
Arg Thr Thr Thr Met Lys Gln Val Leu Phe Tyr Leu Gln Tyr Gln Ile
50 55 60
ATG ACT AAA CGA TTA TAT GAT GAG AAG CAA CAA CAT ATT GTA TAT TGT 542
Met Thr Lys Arg Leu Tyr Asp Gln Lys Gln Gln His Ile Val Tyr Cys
65 70 75
TCA AAT GAT CTT CTA GGA GAT TTG TTT GGC GTG CCA AGC TTC TCT GTG 590
Ser Asn Asp Leu Leu Gly Asp Leu Phe Gly Val Pro Ser Phe Ser Val
80 85 90
AAA GAG CAC AGG AAA ATA TAT ACC ATG ATC TAC AGG AAC TTG GTA GTA 638
Lys Gln His Arg Lys Ile Tyr Thr Met Ile Tyr Arg Asn Leu Val Val
95 100 105
GTC AAT CAG CAG GAA TCA TCG GAC TCA GGT ACA TCT GTG AGT GAG AAC 686
Val Asn Gln Gln Gln Ser Ser Asp Ser Gln Thr Ser Val Ser Gln Asn
110 115 120 125

```

-continued-

AGG	TGT	CAC	CTT	GAA	GGT	GCG	AGT	GAT	CAA	AAG	GAC	CTT	GTA	GAA	GAG	714
Arg	Cys	His	Leu	Glu	Gly	Gly	Ser	Asp	Gln	Lys	Asp	Leu	Val	Gln	Gln	
				130					135					140		
CTT	CAG	GAA	GAG	AAA	GCT	TCA	TCT	TCA	CAT	TTG	ATT	TCT	AGA	CCA	TCT	762
Leu	Glu	Glu	Glu	Lys	Pro	Ser	Ser	Ser	His	Leu	Ser	Arg	Pro	Ser		
				145					150							
ACC	TCA	TCT	AGA	AGG	AGA	GCA	ATT	AGT	GAG	ACA	GAA	GAA	AAT	TCA	GAT	810
Thr	Ser	Ser	Arg	Arg	Arg	Ala	Ile	Ser	Glu	Thr	Glu	Ala	Ser	Asp		
				160					165							
GAA	TTA	TCT	GGT	GAA	GGA	CAA	AGA	AAA	CGC	ACA	AAA	TCT	GAT	AGT	ATT	878
Ser	Leu	Ser	Gly	Glu	Arg	Gln	Arg	Lys	His	Lys	Ser	Asp	Ser	Ser	Ile	
				155					185							
TCC	CTT	TCC	TTT	GAT	GAA	AGC	CTG	GCT	CTG	TGT	GTA	ATA	AGG	GAG	ATA	924
Ser	Leu	Ser	Phe	Asp	Glu	Ser	Leu	Ala	Leu	Cys	Val	Ile	Arg	Glu	Ile	
				180					200					205		
TGT	TGT	CAA	AGA	AGC	AGT	AAC	AGT	GAA	TCT	ACA	GGG	ACG	CCA	TCC	AAT	974
Cys	Cys	Glu	Arg	Ser	Ser	Ser	Ser	Glu	Ser	Thr	Gly	Thr	Phe	Ser	Ala	
				210					215					220		
CCG	GAT	CTT	GAT	GCT	GGT	GTA	AGT	GAA	CAT	TCA	GGT	GAT	TGT	TTG	GAT	1022
Pro	Ala	Leu	Asp	Ala	Gly	Val	Ser	Glu	His	Ser	Gly	Asp	Trp	Leu	Asp	
				275					280					285		
CAG	GAT	TCA	GTT	TCA	GAT	CAG	TTT	AGT	GTA	GAA	TTT	GAA	GTT	GAA	TCT	1070
Glu	Asp	Ser	Val	Ser	Asp	Glu	Phe	Ser	Val	Glu	Phe	Glu	Val	Glu	Ser	
				240					245					250		
CTC	GAC	TCA	GAA	GAT	TAT	AAC	CTT	AGT	GAA	GAA	GGA	CAA	GAA	CTC	TCA	1118
Leu	Asp	Ser	Glu	Asp	Tyr	Ser	Leu	Ser	Glu	Glu	Gly	Glu	Ala	Leu	Ser	
				255					260					265		
GAT	GAA	GAT	GAT	GAG	GTA	TAT	CAA	GTT	ACT	GTG	TAT	CAG	GCA	GGG	GAG	1166
Asp	Glu	Asp	Asp	Glu	Val	Tyr	Gln	Val	Thr	Val	Tyr	Glu	Ala	Gly	Glu	
				270					275					285		
AGT	GAT	ACA	GAT	TCA	TTT	GAA	GAA	GAT	CCT	GAA	ATT	TCC	TTA	GCT	GAC	1214
Ser	Asp	Thr	Asp	Ser	Phe	Glu	Glu	Ala	Pro	Glu	Ile	Ser	Leu	Ala	Asp	
				290					295					300		
TAT	TGG	AAA	TGC	ACT	TCA	TGC	AAT	GAA	ATG	AAI	CCC	CCC	CTT	CCA	TCA	1262
Tyr	Trp	Lys	Cys	Thr	Ser	Cys	Asn	Gln	Met	Asn	Pro	Pro	Leu	Pro	Ser	
				305					310					315		
CAT	TGC	AAC	AAA	TGT	TGG	GCC	CTT	CGT	GAG	AAT	TGG	CTT	CCT	GAA	GAT	1310
His	Cys	Asa	Arg	Cys	Trp	Ala	Leu	Arg	Glu	Asn	Trp	Leu	Pro	Glu	Asp	
				320					325					330		
AAA	GGG	AAA	GAT	AAA	GGG	GAA	ATC	TCT	GAG	AAA	GCC	AAA	CTG	GAA	AAC	1358
Lys	Gly	Lys	Asp	Lys	Gly	Glu	Ile	Ser	Glu	Lys	Ala	Lys	Leu	Glu	Asn	
				325					340					345		
TCA	ACA	CAA	GCT	GAA	GAG	GGC	TTT	GAT	ATT	CCT	GAT	TGT	AAA	AAA	ACT	1406
Ser	Thr	Glu	Ala	Glu	Glu	Gly	Phe	Asp	Val	Pro	Asp	Cys	Lys	Lys	Thr	
				350					355					360		
ATA	GTG	AAT	GAT	TCC	AGA	GAG	TCA	TGT	GTT	GAG	GAA	AAT	GAT	GAT	AAA	1454
Ile	Val	Asn	Asp	Ser	Arg	Gln	Ser	Cys	Val	Glu	Glu	Asn	Asp	Asp	Lys	
				370					375					380		
ATT	ACA	CAA	GCT	TCA	CAA	TCA	CAA	GAA	AGT	GAA	GAC	TAT	TCT	CAG	CCA	1502
Ser	Thr	Glu	Ala	Ser	Glu	Ser	Glu	Glu	Ser	Glu	Asp	Tyr	Ser	Gln	Pro	
				385					390					395		
TCA	ACT	TCT	AGT	AGC	ATT	ATT	TAT	AGC	AGC	CAA	GAA	GAT	GTG	AAA	GAG	1550
Ser	Thr	Ser	Ser	Ser	Ile	Ile	Tyr	Ser	Ser	Gln	Glu	Asp	Val	Lys	Glu	
				400					405					410		
TTT	GAA	ACG	GAA	GAA	ACC	CAA	GAC	AAA	GAA	GAG	AGT	GTG	GAA	TCT	AGT	1598
Phe	Glu	Arg	Gln	Gln	Thr	Glu	Asp	Lys	Glu	Glu	Ser	Val	Glu	Ser	Ser	
				415					420					425		
TTG	CCC	CTT	AAT	GCC	ATT	GAA	CCT	TGT	GTG	ATT	TGT	CAA	GGT	GGA	CCT	1646
Leu	Pro	Leu	Asn	Ala	Ile	Glu	Pro	Cys	Val	Ile	Cys	Gln	Gly	Arg	Pro	
				430					435					440		
AAA	AAT	GGT	TGC	ATT	GTG	CAT	GCG	AAA	ACA	GGA	CAT	CTT	ATG	GCC	TGC	1694
Lys	Asn	Gly	Cys	Ile	Val	His	Gly	Lys	Thr	Gly	His	Leu	Met	Ala	Cys	

-continued-

450												455												460																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
TTT	ACA	TOT	GCA	AAG	AAG	CTA	AAG	AAA	AGG	AAT	AAG	CCC	TGC	CCA	GTA																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						

(2) INFORMATION FOR SEQ ID NO2:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 491 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) SEQUENCE DESCRIPTION: SEQ ID NO2:

Met Cys Asn Thr Asn Met Ser Val Pro Thr Asp Gly Ala Val Thr Thr 1 5 10	
Ser Glu Ile Pro Ala Ser Glu Glu Glu Thr Leu Val Arg Pro Lys Pro 20 25 30	
Leu Leu Leu Lys Leu Leu Lys Ser Val Gly Ala Glu Lys Asp Thr Tyr 35 40 45	
Thr Met Lys Glu Val Leu Phe Tyr Leu Gly Glu Tyr Ile Met Thr Lys 50 55 60	
Arg Leu Tyr Asp Glu Lys Glu Glu His Ile Val Tyr Cys Ser Asn Asp 65 70 75 80	
Leu Leu Gly Asp Leu Phe Gly Val Pro Ser Phe Ser Val Lys Glu His 85 90 95	
Arg Lys Ile Tyr Thr Met Ile Tyr Arg Asn Leu Val Val Val Asn Glu 100 105 110	
Gln Glu Ser Ser Asp Ser Gly Thr Ser Val Ser Glu Asn Arg Cys His 115 120 125	
Leu Glu Gly Gly Ser Asp Glu Lys Asp Leu Val Glu Glu Leu Glu Glu 130 135 140	
Glu Lys Pro Ser Ser Ser His Leu Val Ser Arg Pro Ser Thr Ser Ser 145 150 155 160	
Arg Arg Arg Ala Ile Ser Glu Thr Glu Glu Asn Ser Asp Glu Leu Ser 165 170 175	
Gly Glu Arg Glu Arg Lys Arg His Lys Ser Asp Ser Ile Ser Leu Ser 180 185 190	
Phe Asp Glu Ser Leu Ala Leu Cys Val Ile Arg Glu Ile Cys Cys Glu 195 200 205	
Arg Ser Ser Ser Ser Glu Ser Thr Gly Thr Pro Ser Asn Pro Asp Leu	



-continued-

310	215	220
Asp Ala Gly Val Ser Glu His Ser Gly Asp Trp Leu Asp Glu Asp Ser 125 230 235 240		
Val Ser Asp Glu Phe Ser Val Glu Phe Glu Val Glu Ser Leu Asp Ser 245 250 255		
Glu Asp Tyr Ser Leu Ser Glu Glu Gly Glu Glu Leu Ser Asp Glu Asp 260 265 270		
Asp Glu Val Tyr Glu Val Thr Val Tyr Glu Ala Gly Glu Ser Asp Thr 275 280 285		
Asp Ser Phe Glu Glu Asp Pro Glu Ile Ser Leu Ala Asp Tyr Trp Lys 290 295 300		
Cys Thr Ser Cys Asn Glu Met Asn Pro Pro Leu Pro Ser His Cys Asn 305 310 315 320		
Arg Cys Trp Ala Leu Asp Glu Asn Trp Leu Pro Glu Asp Lys Gly Lys 325 330 335		
Asn Lys Gly Glu Ile Ser Glu Lys Ala Lys Leu Glu Asn Ser Thr Glu 340 345 350		
Ala Glu Glu Gly Phe Asp Val Pro Asp Cys Lys Lys Thr Ile Val Asn 355 360 365		
Asp Ser Arg Glu Ser Cys Val Glu Glu Asn Asp Asp Lys Ile Thr Glu 370 375 380		
Ala Ser Glu Ser Glu Glu Ser Glu Asp Tyr Ser Glu Pro Ser Thr Ser 385 390 395 400		
Ser Ser Ile Ile Tyr Ser Ser Glu Glu Asp Val Lys Glu Phe Glu Arg 405 410 415		
Glu Glu Thr Glu Asp Lys Glu Glu Ser Val Glu Ser Ser Leu Pro Leu 420 425 430		
Asn Ala Ile Glu Pro Cys Val Ile Cys Glu Gly Arg Pro Lys Asn Gly 435 440 445		
Cys Ile Val His Gly Lys Thr Gly His Leu Met Ala Cys Phe Thr Cys 450 455 460		
Ala Lys Lys Leu Lys Lys Arg Asn Lys Pro Cys Pro Val Cys Arg Glu 465 470 475 480		
Pro Ile Glu Met Ile Val Leu Thr Tyr Phe Pro 485 490		

## (2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 1710 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) ORIGINAL SOURCE:

(A) ORGANISM: Mus musculus

(ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 202..1668

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAGGAGCCGC CACCCTTCG TCGCTCGAAG TCTGGACGAC CATGGTCGGT CAGGCCCCCT 60  
 CCGCGGGGCC TCGCGCTCC CCGTGAAGGG TCGGAAGATG CCGGGGAAGT ACGACCGCTC 120  
 TCGTGGGCGA GCGGGAGACC GACCGGACAC CCGTGGGGGA CCTCTCGGAA TCACCGCGCT 180

-continued-

TCTCTCTGGG	CCTCCAGGCC	A	ATG	TGC	AAT	ACC	AAC	ATG	TCT	GTG	TCT	ACC	231			
			Met	Cys	Asn	Thr	Asn	Met	Ser	Val	Ser	Thr				
			1				5					10				
GAG	GGT	GCT	GCA	AGC	ACC	TCA	CAG	ATT	CCA	GCT	TGG	GAA	CAA	GAG	ACT	279
Glu	Gly	Ala	Ala	Ser	Thr	Ser	Glu	Ile	Pro	Ala	Ser	Glu	Glu	Glu	Thr	
				13					20					23		
CTG	GTT	AGA	CCA	AAA	CCA	TTG	CTT	TTG	AAG	TTG	TTA	AAG	TCC	GTT	GGA	327
Leu	Val	Arg	Pro	Lys	Pro	Leu	Leu	Leu	Lys	Leu	Leu	Lys	Ser	Val	Gly	
			30				35					40				
GCG	CAA	AAC	GAC	ACT	TAC	ACT	ATG	AAA	GAG	ATT	ATA	TTT	TAT	ATT	GGC	375
Ala	Glu	Asn	Asp	Thr	Tyr	Thr	Met	Lys	Glu	Ile	Ile	Phe	Tyr	Ile	Gly	
			45				50					55				
CAG	TAT	ATT	ATG	ACT	AAG	AGG	TTA	TAT	GAC	GAG	AAG	CAG	CAG	CAC	ATT	423
Gln	Tyr	Ile	Met	Thr	Lys	Arg	Leu	Tyr	Asp	Glu	Lys	Gln	Gln	His	Ile	
			60			65					70					
GTG	TAT	TGT	TCA	AAT	GAT	CTC	CTA	GGA	GAT	GTG	TTT	GGA	GTC	CCG	AGT	471
Val	Tyr	Cys	Ser	Asn	Asp	Leu	Leu	Gly	Asp	Val	Phe	Gly	Val	Pro	Ser	
					80					85				90		
TTG	TCT	GTG	AAO	GAG	CAC	AGG	AAA	ATA	TAT	GCA	ATG	ATC	TAC	AGA	AAT	519
Phe	Ser	Val	Ile	Glu	His	Arg	Lys	Ile	Tyr	Ala	Met	Ile	Tyr	Arg	Asn	
				95				100						105		
TTA	GTG	GCT	ATA	AOT	CAG	CAA	GAC	TCT	GCG	ACA	TGC	CTG	AGT	GAG	AGC	567
Leu	Val	Ala	Val	Ser	Gln	Gln	Asp	Ser	Gly	Thr	Ser	Leu	Ser	Glu	Ser	
				110				115					120			
AGA	CCT	CAG	CCT	GAA	GGT	GGG	AOT	GAT	CTG	AAG	GAT	CCT	TTG	CAA	GCG	615
Arg	Arg	Gln	Pro	Glu	Gly	Gly	Ser	Asp	Leu	Lys	Asp	Pro	Leu	Gln	Ala	
				125				130					135			
CCA	CCA	GAA	GAG	AAA	CCT	TCA	TCT	TCT	GAT	TTA	ATT	TCT	AGA	CTG	TCT	663
Pro	Pro	Gln	Gln	Lys	Pro	Ser	Ser	Ser	Asp	Leu	Ile	Ser	Arg	Leu	Ser	
				140		145					150					
ACC	TCA	TCT	AGA	AGG	AGA	TCC	ATT	AGT	GAG	ACA	GAA	GAG	AAC	ACA	GAT	711
Thr	Ser	Ser	Arg	Arg	Arg	Ser	Ile	Ser	Glu	Thr	Glu	Glu	Asn	Thr	Asp	
						160				165				170		
GAG	CTA	CCT	GGG	GAG	CGG	CAC	CGG	AAO	CGC	CGC	AGG	TCC	CTG	TCC	TTT	759
Glu	Leu	Pro	Gly	Glu	Arg	His	Arg	Lys	Arg	Arg	Arg	Ser	Leu	Ser	Phe	
				175				180					185			
GAT	CCG	AGC	CTG	GGT	CTG	TGT	GAG	CTG	AGG	GAG	ATG	TGC	AGC	GGC	GGC	807
Asp	Pro	Ser	Leu	Gly	Leu	Cys	Gln	Leu	Arg	Glu	Met	Cys	Ser	Gly	Gly	
				190				195					200			
ACG	AGC	AGC	AGT	AGC	AGC	AGC	AGC	AGC	GAG	TCC	ACA	GAG	ACG	CCC	TCG	855
Thr	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Gln	Ser	Thr	Glu	Thr	Pro	Ser	
								210				215				
CAT	CAG	GAT	CTT	GAC	GAT	GGC	GTA	AGT	GAG	CAT	TCT	GGT	GAT	TGC	CTG	903
Ris	Gln	Asp	Leu	Asp	Asp	Asp	Val	Ser	Gln	His	Ser	Gly	Asp	Cys	Leu	
						225						230				
GAT	CAG	GAT	TCA	GTT	TCT	GAT	CAO	TTT	ABC	GTG	GAA	TTT	GAA	GTT	GAG	951
Asp	Gln	Asp	Ser	Val	Ser	Asp	Gln	Phe	Ser	Val	Glu	Phe	Gln	Val	Glu	
						240				245				250		
TCT	CTG	AGC	TGC	GAA	GAT	TAC	AGC	CTG	AGT	GAC	GAA	GGG	CAC	GAG	CTC	999
Ser	Leu	Asp	Ser	Glu	Asp	Tyr	Ser	Leu	Ser	Asp	Gln	Gly	His	Glu	Leu	
				255				260				265				
TCA	GAT	GAG	GAT	GAT	GAG	GTC	TAT	CGG	GTC	ACA	GTC	TAT	CAG	ACA	GGA	1047
Ser	Asp	Gln	Asp	Asp	Gln	Val	Tyr	Arg	Val	Thr	Val	Tyr	Gln	Thr	Gly	
				270				275					280			
GAA	AGC	GAT	ACA	GAC	TCT	TTT	GAA	GGA	GAT	CCT	GAG	ATT	TCC	TTA	GCT	1095
Glu	Ser	Asp	Thr	Asp	Ser	Phe	Glu	Gly	Asp	Pro	Glu	Ile	Ser	Leu	Ala	
				285			290					295				
GAC	TAT	TGG	AAO	TGT	ACC	TCA	TGC	AAT	GAA	ATG	AAT	CCT	CCC	CTT	CCA	1143
Asp	Tyr	Trp	Lys	Cyt	Thr	Ser	Cys	Asn	Glu	Met	Asn	Pro	Pro	Leu	Pro	
				300		305					310					

-continued

TCA	CAC	TGC	AAA	AGA	TGC	TGG	ACC	CTT	CGT	GAG	AAC	TGG	CTT	CCA	GAC	1191
Ser	His	Cys	Lys	Arg	Cys	Trp	Thr	Leu	Arg	Glu	Asn	Trp	Leu	Pro	Asp	
315				320					325					330		
GAT	AAG	GGG	AAA	GAT	AAA	GTG	GAA	ATC	TCT	GAA	AAA	GCC	AAA	CTG	GAA	1239
Asp	Lys	Gly	Lys	Asp	Lys	Val	Glu	Ile	Ser	Glu	Lys	Ala	Lys	Leu	Glu	
				325				340						345		
AAC	TCA	GCT	CAG	GCA	GAA	GAA	GCG	TTC	GAT	GTG	CCT	GAT	GGC	AAA	AAG	1287
Asn	Ser	Ala	Glu	Ala	Glu	Glu	Gly	Leu	Asp	Val	Pro	Asp	Gly	Lys	Lys	
				350				355					360			
CTG	ACA	GAG	AAT	GAT	GCT	AAA	GAG	CCA	TGT	GCT	GAG	GAG	GAC	AGC	GAG	1335
Leu	Thr	Glu	Asn	Asp	Ala	Lys	Glu	Pro	Cys	Ala	Glu	Glu	Asp	Tyr	Ser	
				365				370					375			
GAG	AAG	GCC	GAA	CAG	ACG	CCC	CTG	TCC	CAG	GAG	AGT	GAC	GAC	TAT	TCC	1383
Glu	Lys	Ala	Glu	Thr	Pro	Leu	Ser	Glu	Glu	Ser	Asp	Asp	Tyr	Ser		
				380			385						390			
CAA	CCA	TCG	ACT	TCC	AGC	AGC	ATT	GTT	TAT	AGC	AGC	CAA	GAA	AGC	GTO	1431
Gln	Pro	Ser	Thr	Ser	Ser	Ser	Ile	Val	Tyr	Ser	Ser	Glu	Glu	Ser	Val	
				395			400						405		410	
AAA	GAG	TTC	AAG	CAG	GAA	ACG	CAG	CAC	AAA	GAC	GAG	AGT	GTG	GAA	TCT	1479
Lys	Glu	Leu	Lys	Glu	Glu	Thr	Glu	His	Lys	Asp	Glu	Ser	Val	Glu	Ser	
				415				420					425			
AGC	TTC	TCC	CTG	AAT	CCC	ATC	GAA	CCA	TGT	GTG	ATC	TGC	CAA	GGG	CGG	1527
Ser	Phe	Ser	Leu	Asn	Ala	Ile	Glu	Pro	Cys	Val	Ile	Cys	Glu	Gly	Arg	
				430				435					440			
CCT	AAA	AAT	GGC	TGC	ATT	GTT	CAC	GGC	AAG	ACT	GGA	CAC	CTC	ATG	TCA	1575
Pro	Lys	Asn	Gly	Cys	Ile	Val	His	Gly	Lys	Thr	Gly	His	Leu	Met	Ser	
				445			450					455				
TGT	TTC	ACG	TGT	GCA	AAG	AAG	CTA	AAA	AAA	AGA	AAC	AAG	CCC	TGC	CCA	1623
Cys	Phe	Thr	Cys	Ala	Lys	Lys	Leu	Lys	Lys	Arg	Asn	Lys	Pro	Cys	Pro	
				460			465					470				
GTO	TGC	AGA	CAG	CCA	ATC	CAA	ATG	ATT	GTG	CTA	AGT	TAC	TTC	AAC		1668
Val	Cys	Arg	Gln	Pro	Ile	Glu	Met	Ile	Val	Leu	Ser	Tyr	Phe	Asn		
				475			480					485				
TAGCTGACCT GCTCACAAAA ATAGAATTTT ATATTTCCTAA CT																1710

(2) INFORMATION FOR SEQ ID NO4:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 489 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: protein

(14) SEQUENCE DESCRIPTION: SEQ ID NO4:

```

Met Cys Asn Thr Asn Met Ser Val Ser Thr Thr Glu Gly Ala Ala Ser Thr
1      5      10      15
Ser Glu Ile Pro Ala Ser Glu Gln Gly Thr Leu Val Arg Pro Lys Pro
20      25      30
Leu Leu Leu Lys Leu Leu Lys Ser Val Gly Ala Glu Asn Asp Thr Tyr
35      40      45
Thr Met Lys Glu Ile Ile Phe Tyr Ile Gly Gln Tyr Ile Met Thr Lys
50      55      60
Arg Leu Tyr Asp Glu Lys Gln Glu His Ile Val Tyr Cys Ser Asn Asp
65      70      75      80
Leu Leu Gly Asp Val Phe Gly Val Pro Ser Phe Ser Val Lys Glu His
85      90      95
Arg Lys Ile Tyr Ala Met Ile Tyr Arg Asn Leu Val Ala Val Ser Glu
100      105      110
Glu Asp Ser Gly Thr Ser Leu Ser Glu Ser Arg Arg Glu Pro Glu Gly
115      120      125

```

-continued-

Gly	Ser	Asp	Leu	Lys	Asp	Pro	Leu	Gln	Ala	Pro	Pro	Glu	Glu	Lys	Pro
130						135				140					
Ser	Ser	Ser	Asp	Leu	Ile	Ser	Arg	Leu	Ser	Thr	Ser	Ser	Arg	Arg	Arg
145					150					155					160
Ser	Ile	Ser	Glu	Thr	Glu	Glu	Asn	Thr	Asp	Glu	Leu	Pro	Gly	Glu	Arg
			165						170					175	
His	Arg	Lys	Arg	Arg	Ser	Leu	Ser	Phe	Asp	Pro	Ser	Leu	Gly	Leu	
			180				185						190		
Cys	Glu	Leu	Arg	Glu	Met	Cys	Ser	Gly	Gly	Thr	Ser	Ser	Ser	Ser	Ser
		195				200					205				
Ser	Ser	Ser	Glu	Ser	Thr	Glu	Thr	Pro	Ser	His	Gln	Asp	Leu	Asp	Asp
		210				215					220				
Gly	Val	Ser	Glu	His	Ser	Gly	Asp	Cys	Leu	Asp	Gln	Asp	Ser	Val	Ser
			230						235					240	
Asp	Glu	Phe	Ser	Val	Glu	Phe	Glu	Val	Glu	Ser	Leu	Asp	Ser	Glu	Asp
			245					250					255		
Tyr	Ser	Leu	Ser	Asp	Glu	Gly	His	Glu	Leu	Ser	Asp	Glu	Asp	Asp	Glu
		260						265				270			
Val	Tyr	Arg	Val	Thr	Val	Tyr	Glu	Thr	Gly	Glu	Ser	Asp	Thr	Asp	Ser
		275					280					285			
Phe	Glu	Gly	Asp	Pro	Glu	Ile	Ser	Leu	Ala	Asp	Tyr	Trp	Lys	Cys	Thr
			290			295					300				
Ser	Cys	Asn	Gln	Met	Asn	Pro	Pro	Leu	Pro	Ser	His	Cys	Lys	Arg	Cys
			310							315				320	
Trp	Thr	Leu	Arg	Glu	Asn	Trp	Leu	Pro	Asp	Asp	Lys	Gly	Lys	Asp	Lys
			325					330					335		
Val	Glu	Ile	Ser	Glu	Lys	Ala	Lys	Leu	Glu	Asn	Ser	Ala	Gln	Ala	Glu
			340					345					350		
Glu	Gly	Leu	Asp	Val	Pro	Asp	Gly	Lys	Lys	Leu	Thr	Glu	Asn	Asp	Ala
		355				360					365				
Lys	Glu	Pro	Cys	Ala	Glu	Glu	Asp	Ser	Glu	Glu	Lys	Ala	Glu	Gln	Thr
		370				375					380				
Pro	Leu	Ser	Gln	Glu	Ser	Asp	Asp	Tyr	Ser	Glu	Pro	Ser	Thr	Ser	Ser
		385				390				395					
Ser	Ile	Val	Tyr	Ser	Ser	Gln	Glu	Ser	Val	Lys	Glu	Leu	Lys	Glu	Glu
			405						410				415		
Thr	Gln	His	Lys	Asp	Glu	Ser	Val	Glu	Ser	Ser	Phe	Ser	Leu	Asn	Ala
		420					425						430		
Ile	Glu	Pro	Cys	Val	Ile	Cys	Glu	Gly	Arg	Pro	Lys	Asn	Gly	Cys	Ile
		435					440					445			
Val	His	Gly	Lys	Thr	Gly	His	Leu	Met	Ser	Cys	Phe	Thr	Cys	Ala	Lys
		450				455					460				
Lys	Leu	Lys	Lys	Arg	Ala	Lys	Pro	Cys	Pro	Val	Cys	Arg	Glu	Pro	Ile
		465				470				475					480
Glu	Met	Ile	Val	Leu	Ser	Tyr	Phe	Asn							
			485												

We claim:

1. A method of screening for a neoplastic tissue in a human comprising:

detecting amplification of a human MDM2 gene or elevated expression of a human MDM2 gene by detecting human MDM2 mRNA in a tissue or body fluid isolated from a human, wherein amplification of the human MDM2 gene or elevated expression of the human MDM2 gene provides a method of

screening for neoplasia or the potential for neoplastic development.

2. The method of claim 1 wherein gene amplification is detected.

3. The method of claim 1 wherein said mRNA is detected by Northern blot analysis by hybridizing mRNA from said tissue to a human MDM2 nucleotide probe.

4. The method of claim 3 wherein the human MDM2 nucleotide probe comprises nucleotides 1-2372 of

human MDM2, as shown in FIG. 1, or fragments thereof consisting of at least 14 contiguous nucleotides.

5. The method of claim 2 wherein the gene amplification is detected using polymerase chain reaction.

6. The method of claim 2 wherein amplification of the human MDM2 gene is detected by Southern blot analysis wherein the human MDM2 gene is hybridized with a nucleotide probe which is complementary to hMDM2 DNA.

7. The method of claim 2 wherein gene amplification is determined by comparing the copy number of hMDM2 in the tissue to the copy number of hMDM2 in a normal tissue of the human.

8. The method of claim 1 wherein elevated expression of a human MDM2 gene is determined by comparing the amount of hMDM2 mRNA in the tissue to the

amount of hMDM2 mRNA in a normal tissue of the human.

9. The method of claim 2 wherein gene amplification is detected when at least 3-fold more hMDM-2 DNA is observed in the tissue relative to a control sample comprising a normal tissue.

10. The method of claim 1 wherein elevated expression is detected when at least 3-fold more hMDM-2 mRNA is observed in the tissue relative to a control sample comprising a normal tissue.

11. The method of claim 1 wherein the neoplasia is a sarcoma.

12. The method of claim 11 wherein the sarcoma is a liposarcoma, malignant fibrous histiocytoma, or osteosarcoma.

\* \* \* \* \*

20

25

30

35

40

45

50

55

60

65

UNITED STATES PATENT AND TRADEMARK OFFICE  
**CERTIFICATE OF CORRECTION**

PATENT NO. : 5,411,860

DATED : May 2, 1995

INVENTOR(S) : Bert Vogelstein and Kenneth Kinzler

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

At column 1, line 3 insert the following:

"This invention was made with government support under  
Grant CA43460, CA35494 and CA41183 awarded by the  
National Institutes of Health. The government has certain rights  
in this invention."

Signed and Sealed this  
Fifth Day of August, 1997



Attest:

Attesting Officer

BRUCE LEHMAN

Commissioner of Patents and Trademarks

010-44110  
BEST AVAILABLE COPY

SECOND EDITION

# Recombinant DNA



James D. Watson  
Michael Gilman

Jan Witkowski  
Mark Zoller

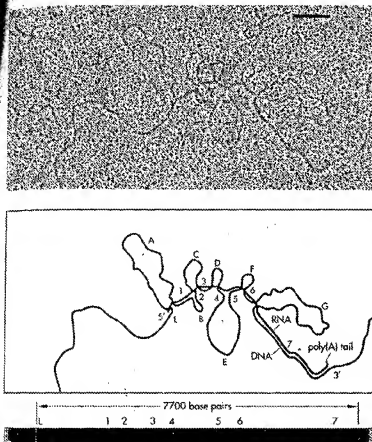


FIGURE 8-2

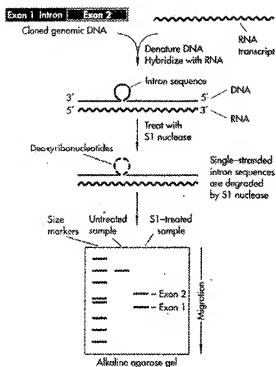
Examination of gene structure by electron microscopy of DNA-mRNA hybrids. DNA containing the gene for ovalbumin was hybridized with ovalbumin mRNA. The regions of the gene that hybridize to the mRNA are eight exons (1-7). Genomic DNA that encodes introns does not hybridize to the mRNA but forms seven loops (A-G). The upper portion of the figure shows the actual electron micrograph. Regions where genomic DNA hybridized to mRNA form a thicker line than do the single-stranded genomic DNA loops. The locations and lengths of the introns were estimated simply by plotting the position of each loop along the mRNA molecule. Because the length of the mRNA (in nucleotides) was known, the approximate positions of the introns could be calculated. The middle portion of the figure shows the interpretation of the electron micrograph. The 5' and 3' ends of the mRNA are indicated. The bottom portion shows a scheme of the structure of the ovalbumin gene determined subsequently by DNA sequence analysis of the exon-intron boundaries in genomic DNA and comparison with the cDNA sequence. The exons are shown in green and the introns are shown in red (modified from Chambon, 1981).

### Introns Are Discovered in Eukaryotic Genes

Excited by these results, people working on the structure of eukaryotic genes searched for splicing of cellular RNAs. Within a very short time after the discovery of adenovirus splicing, the coding sequences of  $\beta$ -globin, ovalbumin, and immunoglobulin genes were also found to be interrupted by noncoding DNA. Proof that chromosomal genes were spliced came initially by electron microscopy. The sizes and locations of these introns were then estimated by a technique called *S1 nuclease mapping* (Figure 8-3). The regions

of the chromosomal DNA not present in the mature mRNA were given the name *introns*. The coding sequences were called *exons*, because the processed mRNAs, without the introns, "exit" from the nucleus to the cytoplasm. It should be noted that at the time that the electron microscopy experiments on adenovirus were done, no one had cloned a cellular gene yet. Once the first genes were cloned, introns were identified by comparing the cloned genomic DNA with the corresponding cloned cDNA. For small genes, such as the  $\beta$ -globin gene, the sizes of the introns and the locations of intron-exon boundaries were precisely determined by sequencing cloned genomic DNA and





**FIGURE 8-3**  
S1 nuclease mapping finds introns in a gene. Total cellular RNA is hybridized to a cloned genomic DNA fragment that contains a single intron. An RNA-DNA heteroduplex is formed between the gene DNA and its corresponding mRNA by base-pairing of complementary sequences. Intron sequences in the gene do not hybridize with the mRNA, so these DNA sequences form single-stranded loops. The RNA-DNA hybrid is treated with S1 nuclease, an enzyme that digests single-stranded DNA into mononucleotides. The unpaired intron sequences thus are digested, a process that splits the genomic DNA into two fragments. The sizes of these two DNA fragments are determined by gel electrophoresis. From this information, the length and location of the intron in the gene are deduced.

comparing the sequence with the cDNA sequence and with the protein sequence.

Introns exist in genes from all eukaryotic animals, in plant genes, and, surprisingly, in genes of the *E. coli* phage T4. Often the introns of a gene contain many more nucleotides than do its coding exons, thus ac-

counting for the previously unexplained large sizes of many primary RNA transcripts. The number and size of introns vary widely from one gene to another. Two introns are present in all genes of the  $\beta$ -globin family (Figure 8-4). The sizes of introns in the  $\beta$ -globin genes from different species differ slightly, but their positions are always the same relative to the coding sequence. The ovalbumin gene is more complicated than the  $\beta$ -globin gene and contains seven introns. The length of introns can vary from 31 nucleotides in an SV40 gene to over 210,000 nucleotides in the human dystrophin gene. A few genes, such as the genes coding for the  $\alpha$  and  $\beta$  forms of interferon and most of the genes from the yeast *Saccharomyces cerevisiae* do not contain introns. Intronless mammalian genes can also be generated through recombinant DNA tricks and tested to see how they function in vivo. For certain genes, the complete removal of introns has no consequence. Such genes produce fully active mRNA transcripts. However, for other genes, the removal of their natural introns somehow blocks the exit of functional mRNA products to the cytoplasm. Perhaps in these latter cases the newly made transcripts adopt configurations incompatible with exit from the nucleus.

### Specific Base Sequences Are Found at Exon-Intron Boundaries

By the summer of 1978, just a year after the first split genes were discovered, the sequences of many exon-intron boundaries from cellular genes had been determined. It was hoped that such sequence data would be useful in predicting the location of an intron in a gene and would also explain how splicing was accomplished. People expected to find that the sequence at the upstream (5') and downstream (3') ends of an intron would be complementary. These sequences would therefore be expected to hybridize and form a stretch of double-stranded RNA, which would be recognized and precisely excised by specific splicing enzymes. However, this idea was soon discounted, since the sequences at the ends of the introns were not complementary. The upstream and downstream splice sites therefore could not be brought together by self-



Search

mdm2 and gei

Return to Results | Remove From



Terms of Use



GeneCards Homepage



Search Examples



Contact Us

GeneCard for gene **MDM2**  
GC12P067488

Approved [UCL/HGNC/HUGO Human Gene Nomenclature database](#) symbol  
**MDM2** (Mdm2, transformed 3T3 cell double minute 2, p53 binding protein (mouse))

**Aliases and Descriptions**  
(According to [GDB](#) [OMIM](#)  
[HUGO](#) [LocusLink](#)  
[SWISS-PROT](#) [TrEMBL](#)  
and/or [GeneLoc](#))

- hdm2 (u)
- Mdm2, transformed 3T3 cell double minute 2, p53 binding protein (mouse) (cos. 1)
- mouse double minute 2, human homolog of, p53-binding protein (see)
- p53-binding protein MDM2 (u)
- Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein Mdm2) (Oncop

Previous GC identifiers: GC12M086977 GC12P069000 GC12P069018

**Chromosomal Location**  
(According to [GeneLoc](#) and/or  
[HUGO](#) and/or [LocusLink](#)  
(NCBI build 34)  
Genomic Views According to [UCSC](#)  
and [Ensembl](#))

**Chromosome: 12** [GeneLoc gene densities](#)

[LocusLink](#) cytogenetic band: **12q14.3-q15** [Ensembl](#) cytogenetic band: **12q15**

**Gene in genomic location: bands according to Ensembl, locations according to Gene**



**GeneLoc location for GC12P067488:** (about GC identifiers)

**Start:** 67,488,238 bp from pter  
**End:** 67,520,481 bp from pter  
**Size:** 32,243 bases  
**Orientation:** plus strand

Genome View

[UCSC Golden Path with GeneCards custom track](#)

**Proteins**  
(According to  
[SWISS-PROT](#) [TrEMBL](#)  
and/or [MIPS](#) PDB rendering  
according to [OCA](#))

Swiss-Prot **MDM2 HUMAN**

- **Size:** 491 amino acids; 55232 Da
- **Cofactor:** ZINC IS REQUIRED FOR UBIQUITIN LIGASE E3 ACTIVITY.
- **Subunit:** BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND SPECIFIC WITH RETINOBLASTOMA PROTEIN (RB), E1A-ASSOCIATED PROTEIN P300 AI
- **Subcellular location:** NUCLEAR AND CYTOPLASMIC. EXPRESSED PREDOMINANTLY IN NUCLEOLAR LOCALIZATION SIGNALS IN BOTH ARF(P14) AND MDM2 MAY BE NUCLEOLAR LOCALIZATION OF BOTH PROTEINS.
- **Alternative products:** Alternative splicing.
- **Post-translational modifications:** PHOSPHORYLATED IN RESPONSE TO IONIZING RADIATION.
- **Miscellaneous:** MDM2 RING FINGER MUTATIONS THAT FAILED TO UBIQUITINATE P53 FOR DEGRADATION WHEN EXPRESSED IN CELLS.
- **3D structure:** PDB id 1YCR (3D)

**MIPS Pedant Viewer:** [74112](#) [74106](#) [74107](#) [74108](#) [74109](#)

**REFSEQ proteins (5 alternative transcripts):** [NP\\_002383.1](#) [NP\\_006869.1](#) [NP\\_006870.1](#) [NP\\_006871.1](#)

**Protein Domains/Families**  
(According to [InterPro](#)  
[ProtoNet](#)  
[SWISS-PROT](#) [TrEMBL](#))

**InterPro Domains and Families:**

- [IPR001876](#) Znf\_Fam50P
- [IPR003121](#) SWB
- [IPR001841](#) Znf\_rmg

## YAC clones that extend the human Chromosome 12cen-12q15 region contig map

S.E. Andersen, N.J. Lench, A.F. Markham

Molecular Medicine Unit, University of Leeds, Clinical Sciences Building, St James's University Hospital, Leeds LS9 7TF, UK

Received: 4 March 1996 / 5 June 1996

Human Chromosome (Chr) 12q13 is a region of clinical interest in that a variety of disease phenotypes have been localized to the area. A number of relatively detailed genetic maps of the region around human Chr 12q13 are now available. A second-generation YAC contig map of human Chr 12, which extensively covers this area, has also been produced. This has resolved a number of discrepancies in the previous genetic maps. The current YAC contig map contains only five small gaps in the approximately 45 cM region between the polymorphic microsatellite markers *D12S333* and *D12S106*. A number of expressed sequences have also been localized with varying degrees of precision to the region around 12q13.

In an attempt to generate YAC resources with low levels of chimerism and to facilitate disease and EST mapping in this genomic area, we now describe the isolation of 130 YACs from the ICI library (Anand et al. 1990). We have used 36 genetic markers and 11 ESTs from five previous Chr 12 maps (Guyer and Cann, 1992; Weissenbach et al. 1992; Schoenmakers et al. 1994; Kucherlapati et al. 1994; Guyapay et al. 1994). Our data can be integrated with the most recent second-generation YAC contig map of Chr 12 (Krauter et al. 1995). We have mapped the *WNT-1* gene (Nusse et al. 1991) at 62–66 cM (based on the 200 cM overall Chr 12 map of Krauter et al. 1995) in close proximity to *D12S339* (in the contig gap between *D12S85* and *D12S361*) and the *GPD1* gene. This is an area under-represented in the latest Chr 12 contig map. We have also observed close linkage between *GADD153* and *GLI* at 77 cM in the interval between *D12S312* and *D12S90* (Krauter et al. 1995). The markers *D12S17* and *D12S96* also co-localize at 71 cM in the area between *D12S390* and *D12S398*, which has enabled us to position several first-generation genetic markers with respect to these points. We have isolated YACs for *LRP2/AMR*, which is known to be physically linked to *GLI* (Forus and Myklesbost 1992). These may be of value in closing a gap at 76 cM in the current map immediately centromeric of *GADD153/GLI*. The YAC 11GH7 containing *D12S312* may also be useful in this respect as it is the closest known marker proximal to the same contig gap. Similarly, YACs containing *MDM2* and *D12S43* may help to close the gap between *RAP1B* and *D12S80* at 85 cM. As well as YACs containing *COL2A1* and *LALBA*, which are included in the latest contig map, we have isolated YACs for *PMCA*, *CD63*, and *PAB1*, which are as yet not accurately positioned on the map. *D12S39* or *IGF-1* containing YACs described herein are now known to map distal to our region of interest at 113 cM and 118 cM respectively (Krauter et al. 1995). Vectorized YAC end sequencing (Riley et al. 1990) has enabled us to generate other novel STSs and corresponding YACs from the region.

Novel YACs identified with 36 genetic markers from the 12q13 region and from 11 expressed sequences in the region are presented in Table 1. Where known, a corresponding YAC from the second-generation YAC contig map of human Chr 12 is in-

cluded in Table 1 for ease of cross-reference (Krauter et al. 1995). The 47 loci screened all produced at least one positive YAC and, in most cases, multiple YACs, consistent with a  $\times 3.5$ -fold genomic representation of this library (Anand et al. 1990). In a number of cases, the same YAC was identified with more than one genetic marker, implying that these markers lie in close physical proximity (the average insert size of the YACs in this library is 350 kb). *D12S39* and the *WNT-1* oncogene both gave positive signals with YACs 26BB4, 28AD4, and 32BC2 in the region of the 64 cM contig gap (Fig. 1). *D12S368* and *D12S174* at around 68 cM both identified YAC 30AD11. *D12S17* and *D12S96* both identified the YACs 20BB4 and 20BF12. This places *D12S17* at the 71 cM locus in the middle of the KRT gene family and centromeric of *D12S398*, *D12S359*, *D12S19*, *D12S325*, *HOXC5*, and *D12S103* respectively. *GADD153* and *GLI1* at 76 cM identify the YAC 26EG10, confirming the close physical linkage of these two genes. This locus is immediately telomeric of a gap at 76 cM in the current second-generation YAC contig map, and end sequencing of 26EG10 may assist in contig closure. The markers *D12S305* and *D12S104* at 77 cM both identify the YAC 10ED1 concordant with the second-generation map. End sequencing of 10ED1 and re-sequencing of the YAC library identified 15BE5. This YAC also contains *D12S90*, confirming marker order *D12S90*, *D12S305*, *D12S104*, *D12S83* and *D12S334* both identify the YAC 34CA7, confirming their physical proximity at 77 cM also (Krauter et al. 1995).

YACs for which sequencing of human DNA insert termini has been performed are shown in Table 1. Following confirmation of localization to Chr 12 with the Coriell Mapping Panel, additional YACs were isolated as shown (Table 1). These increase the YAC coverage at *D12S96/17* (71 cM, 20BF12), *D12S90/305/104* (77 cM, 10ED1, 15BE5), and *D12S43* around the 86 cM region (34HH10). A number of YACs proved to be chimeric during analysis; the right-hand end of 40AA12 (a YAC identified using primers from the *MDM2* gene) maps to human Chr 1. The *MDM2* YAC 6FD11 is also highly chimeric by FISH analysis (data not shown). The *MDM2* YAC 40CB5 was not chimeric by FISH analysis and was used to confirm its localization to the 12q14 region (Fig. 2; Heighway et al. 1994). The other YAC termini sequences in Table 1 all map to Chr 12 by PCR analysis of somatic cell hybrids.

In the last four years, six separate Chr 12 maps have been presented which include markers in and around the 12q13 region. The NIH/CEPH collaborative mapping group comprehensive genetic linkage map of the human genome (Guyer and Cann 1992), in conjunction with data of Schoenmakers and associates (1994), provided an approximate marker order cen-COL2A1-ELA1-D12S79-D12S15-D12S75-D12S14-D12S4-D12S18-D12S16-D12S17-D12S6-D12S28-D12S22-D12S28-D12S19-D12S43-D12S8-D12S64-D12S7-tel, with the latter probably located distally in 12q15. In complementary studies in 1992, Weissenbach and colleagues described another set of markers across the same







Fig. 2. Fluorescent in situ hybridization analysis of the MDM2 YAC 40CB5 confirming localization of this gene to 12q14-15.

questions concerning the frequent translocations encountered in the region in a variety of malignancies.

**Acknowledgments.** Research in the authors' laboratories is supported by the Medical Research Council, the Wellcome Trust, the Yorkshire Cancer Research Campaign and the West Riding Medical Research Trust. We thank the Candelighters for additional provision of specific research support for this project. J.P. Leck for FISH analysis, C.A. Higgins for manuscript preparation and Zeneca plc for library access.

## References

- Anand, R., Riley, J.H., Butler, R., Smith, J.C., Markham, A.F. (1990). A 3.5 genome equivalent multi-access YAC library: construction, characterisation, screening and storage. *Nucleic Acids Res.* 18, 1951-1956.
- Bureau, J.F., Bühl, F., Bralic, M., Le Paslier, D. (1993). The gene coding for interferon- $\gamma$  is linked to the D12S335 and D12S313 microsatellites and to the MDM2 gene. *Genomics* 28, 109.
- Fejzo, M.S., Yoon, S.J., Montgomery, K.T., Rein, M.S., Weremowicz, S., Krauter, K.S., Dorman, T.E., Fletcher, J.A., Mao, J., Moir, D.T., Kucherlapati, R.S., Morton, C.C. (1995). Identification of a YAC spanning the translocation breakpoints in uterine leiomyomata, pulmonary chondroid hamartoma, and Lipomas: physical mapping of the 12q14-q15 breakpoint region in uterine leiomyomata. *Genomics*, 26, 265-271.
- Forus, A., Myklebost, O. (1992). A physical map of a 1.3 Mb region on the long arm of chromosome 12, spanning the GLI and LRP loci. *Genomics*, 14, 117-120.
- Guyer, M., Caan, H. (1992). A comprehensive genetic linkage map of the human genome: NIH/CEPH Collaborative Mapping Group. *Science* (reprint) pp 3-66.
- Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., Weissenbach, J. (1994). The 1993-94 Genethon human genetic linkage map. *Nature Genet.* 7, 246-339.
- Highway, J., Mitchell, E.L.D., Jones, D., White, G.R.M., Santibanez Kofre, M.F. (1994). A transcribed polymorphism and sub-localisation of MDM2. *Hum. Genet.* 93, 611-612.
- Khauf, Z.A., Matsuhashine, H., Valentine, M., Shapiro, D.N., Sherr, C.J., Leck, A.T. (1993). Coamplification of the *CDK4* gene with *MDM2* and *GLI* in human sarcomas. *Cancer Res.*, 53, 5535-5541.
- Krauter, K., Montgomery, K., Yoon, S.J., LeBlanc-Stracessi, J., Renault, B., Marondel, I., Herdman, V., Cupelli, L., Banks, A., Lieman, J., Menninger, J., Bray-Ward, P., Nadkarni, P., Weissenbach, J., LePaslier, D., Rigault, P., Chumakov, I., Cohen, D., Miller, P., Ward, D., Kucherlapati, R. (1995). A second-generation YAC contig map of human chromosome 12. *Nature*, 377, S321-S333.
- Kucherlapati, R., Craig, L., Marynen, P. (1994). Report of the second international workshop on human chromosome 12 mapping 1994. *Cytogenet. Cell Genet.* 67, 245-276.
- Nusse, R., Brown, A., Papkoff, J., Scambler, P., Shackelford, G., McMahon, A., Moon, R., Varnus, H. (1991). A new nomenclature for int-1 and related genes: the Wnt gene family. *Cell*, 64, 231-232.
- Paulsen, S., Sandberg, A.A., Herz, J., Gemmill, R.M. (1992). Putative apolipoprotein receptor gene (LRP, A2MR) is not rearranged in either myxoid liposarcoma or lipomas with translocations in 12q13-14. *Cancer Cent. Cytogenet.*, 60, 125-130.
- Riley, J.H., Butler, R., Ogilvie, D., Finnear, R., Jenner, D.E., Powell, S., Anand, R., Smith, J.C., Markham, A.F. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.*, 18, 2887-2890.
- Schoenmakers, E.F.P.M., Kools, P.F.J., Mols, R., Kazmierczak, B., Bartnitzke, S., Bullerdick, J., Dal Cin, P., De Jong, P.F., Van Den Berghe, H., Van De Ven, W.J.M. (1994). Physical mapping of chromosome 12q breakpoints in lipoma, pleomorphic salivary gland adenoma, uterine leiomyoma and myxoid liposarcoma. *Genomics*, 20, 210-222.
- Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseux, G., Lathrop, M. (1992). A second-generation linkage map of the human genome. *Nature* 359, 794-801.

## SHORT COMMUNICATION

### The Gene Coding for Interferon- $\gamma$ Is Linked to the D12S335 and D12S313 Microsatellites and to the MDM2 Gene

J. F. BUREAU,\* F. BIHL,\* M. BRAHIC,\*<sup>1</sup> AND D. LE PASLIER†

\*Unité des Virus Lents, Institut Pasteur, URA CNRS 1157, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France; and  
†Centre d'Etude du Polymorphisme Humain, 27 rue Juliette Dodu, 75010 Paris, France

Received March 13, 1995; accepted April 27, 1995

Interferon- $\gamma$  is a cytokine with multiple effects. It interferes with the replication of several viruses and plays a key role in the regulation of immune responses. Therefore, the gene coding for interferon- $\gamma$  could be implicated in the susceptibility of humans to several diseases. We have localized this gene close to the D12S335 and D12S313 microsatellites on both the physical and the genetic maps of the human genome. We also physically mapped this gene close to the MDM2 locus on chromosome band 12q15. Finally, we describe the organization of the *IIf*, *Myf-6*, *Mdm1*, and *Mdm2* loci on mouse chromosome 10, in a region syntenic to human chromosome band 12q15. © 1995 Academic Press, Inc.

Interferon- $\gamma$  (IFNG) was discovered because of its antiviral property (12), although it is now studied chiefly because of its central role in the regulation of immune responses. This cytokine is secreted by CD4<sup>+</sup> T cells committed to the Th1 pathway, by CD8<sup>+</sup> T cells, and by activated macrophages. Because of these multiple functions, the gene coding for IFNG is a good candidate gene for control of susceptibility to various infectious as well as immune-mediated diseases of humans and other mammals. The persistent infection of the mouse central nervous system by Theiler's virus is a case in point. The *IIf* gene is a good candidate for control of the persistence of the infection for two reasons: (i) Persistence is controlled by a gene that was mapped to the telomeric region of chromosome 10, close to the *IIf* locus (5). (ii) Resistant 129Sv mice whose gene coding for the IFNG receptor has been inactivated become susceptible (8).

A precise localization of the human IFNG gene was not available until the present work. This gene had been localized to band 12q24 using FISH and by screening a panel of somatic hybrid cell lines (14, 19). Recently, Ruiz-Linares (17) described a microsatellite in the first intron of this gene, and Awata *et al.* (3) reported a difference in the allelic distribution of this marker between a group of patients with insulin-de-

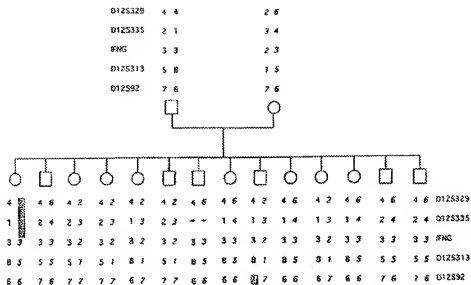
pendent diabetes mellitus and a control group. For future studies using the candidate gene approach, it would be extremely useful to locate precisely the IFNG gene in particular with respect to polymorphic microsatellites markers.

With this goal in mind, we screened the YAC library from CEPH (1) by PCR using two primers (forward 5'-GCTGTTATAATTATAGCTTGT-3' and reverse 5'-AGGGTATTATTATACGAGCT-3') derived from those described by Ruiz-Linares (17) for the IFNG microsatellite. PCR was performed using a kit from Amersham. After denaturation at 94°C for 2 min, 200 ng of DNA in 25  $\mu$ l was submitted to 40 cycles of amplification (94°C, 40 s; 50°C, 40 s; 72°C, 15 s). Using this procedure, we isolated clone 825G7 of the YAC library. We found that clone 825G7 belonged to a contig of 25 other YAC clones, according to inter-*Alu* PCR patterns described in the CEPH-Généthon WWW server (6). All clones in this contig contained either both microsatellites D12S335 and D12S313 or one of the two. Each clone was tested for the presence of the IFNG microsatellite. Eleven clones contained the IFNG microsatellite. Ten also contained both the D12S335 and D12S313 microsatellites (YAC clones 745A10, 743E2, 751A4, 809H4, 823D1, 870H3, 924E4, 926A6, and 983H8), and one, YAC clone 763F1, contained only the D12S313 microsatellite. Therefore, these results demonstrate that the IFNG gene is physically linked to the D12S335 and D12S313 microsatellites.

To confirm the position of the IFNG locus on the genetic map, we analyzed the segregation of alleles of five microsatellite markers in the eight families that have been used to construct the Généthon map (10). Besides markers D12S335 and D12S313, we used two more microsatellites located on either side of the D12S335-D12S313 region. One was taken from a group of six cosegregating markers (D12S104, D12S305, D12S355, D12S334, D12S83, and D12S329) centromeric to the D12S335 marker and the other from a group of three cosegregating markers (D12S344, D12S80, and D12S92) telomeric to the D12S313 marker. For each family, a pair of markers was chosen according to its polymorphism within the family. The parental meiosis were not informative for

<sup>1</sup> To whom correspondence should be addressed. Fax: 33 (1) 40 61 31 67. E-mail: mbrahim@pasteur.fr.

## Family 102



## Family 1331

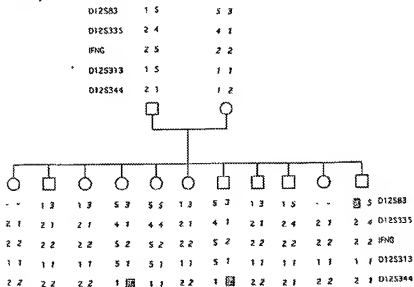


FIG. 1. Segregation of the alleles of IFNG and four microsatellite markers in two families. Maternal alleles are shown italicized. The genotypes of the microsatellites were obtained from the Génethon database. Dashes correspond to cases for which genotypes were not available. Stippled areas indicate a recombination event, and bold face identifies alleles from each parent. The alleles of IFNG are numbered according to the size of the amplified DNA, from the smaller to the larger fragment.

two families. The pedigrees of the other six were analyzed for the markers described above: IFNG, D12S335, D12S313, and the two flanking microsatellites. The results are presented in Fig. 1 for two families and summarized in Table 1. No recombination was observed between the D12S313 and the IFNG markers among 137 informative meioses. Only one recombination was observed between the D12S335 and the IFNG markers among 135

informative meioses. Analysis revealed that the flanking markers were indeed less linked to the IFNG locus than the D12S335 and D12S313 markers.

We next showed, by physical and genetic mapping, that the IFNG locus is linked to the D12S313 and D12S335 loci. Indeed, YAC clone 926A6 was recognized by these three markers. This YAC clone is chimeric and contains regions of both chromosome 12 (band



TABLE 1  
Genetic Localization of IFNG

	No. crossing over with IFNG	No. of informative meioses
D12S329/D12S83	5 (3.9 cM)	129
D12S335	1 (0.8 cM)	135
D12S313	0 (0.0 cM)	137
D12S344/D12S92	6 (4.4 cM)	137

12q15) and chromosome 5, as shown by FISH (6, 9). Therefore, our results localize the IFNG locus to band 12q15, which is different from the previously published localization (band 12q24) (19). The region of human chromosome 12 to which we localize the IFNG gene is syntenic to a region of the mouse genome where we and others have previously localized the mouse *Ilg* gene (2, 5, 18). In our work, the positions of loci *Mdm1*, *Ilg*, and *Myf6* and of three microsatellites D10Mit10, D10Mit14, and D10Mit164 from the Whitehead Institute were determined using the progeny of a F1 (B10.S × SJL/J) × B10.S backcross (Fig. 2) (5). This backcross was typed for more than 100 loci that were distributed among all of the autosomal chromosomes. The results showed that the *Mdm1*, *Ilg*, and *Myf6* loci were linked to each other in the telomeric region of chromosome 10. Other authors have also reported that the *Mdm1*, *Mdm2*, *Mdm3*, and *Ilg* loci are closely linked (2, 18). Interestingly, the MDM2 locus has been localized by FISH to human chromosome 12q14.3–q15 (11). Thus, the analysis of the syntenic region of the mouse genome agrees with our localization of the IFNG locus to human chromosome 12q15.

We confirmed the physical linkage of the MDM2 and IFNG loci by screening, with the MDM2 PCR (11), the contig of YAC clones recognized by at least one of the D12S335, D12S313, and IFNG microsatellites. The MDM2 marker recognized three YAC clones (751A4,

870H3, and 983H8) that were also recognized by the D12S335, D12S313, and IFNG microsatellites. Thus, the MDM2 and IFNG loci are physically linked. The MDM2 gene, which codes for a p53-like protein, is located in a region associated with several cancers (15, 16). New polymorphic markers for this region are of particular interest since the only one available so far is a *Nla*IV polymorphic restriction site (11).

The *Myf6* locus is located in the region of mouse chromosome 10 syntenic to human chromosome band 12q15 (4). Therefore, we decided to localize the MYF-6 locus physically. We screened the CEPH YAC library using MYF-6-specific primers (forward 5'-AGAOCCTTCTCC-ACGCAGCAG-3' and reverse 5'-GCGAAATCTGTG-TGCAGCT-3') under PCR conditions identical to those described above for the MDM2 marker. Two clones, 921C6 and 982A6, were isolated. We found that both belonged to a contig of nine other YAC clones, according to inter-*Alu* PCR patterns described in the CEPH-Génethon WWW server (6). Three of them, clones 940B8, 937D9, and 949H4, contained the MYF-6 marker. According to the CEPH-Génethon server, clones 940B8 and 937D9 contained the D12S106 microsatellite, a marker 15 cM from the D12S313 microsatellite, toward the telomere (10). Figure 3 shows the position of these various markers on human chromosome 12q15 and mouse chromosome 10. The organization of the syntenic regions is very similar. IFNG/D12S313 and D12S106 are 15 cM apart, whereas *Ilg*/Mdm1 and *Myf6* are 5 cM apart. However, since 1 cM is, on average, equivalent to 1.7 Mb in the mouse and 1.0 Mb in human, the physical distances between the markers in mouse and in human could be of the same order.

In conclusion, we have shown that the IFNG and the MDM2 genes are close to each other and to the D12S335 and D12S313 microsatellites. These markers are most likely localized to chromosome bands 12q14.3–q15, in a region syntenic to the telomeric part

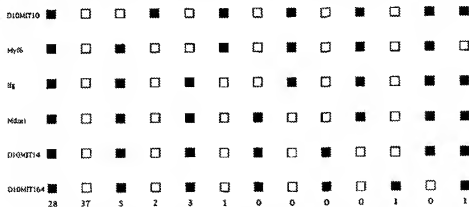


FIG. 2. Segregation of the alleles of *Ilg*, *Mdm1*, *Myf6*, and some markers published by the Whitehead Institute (7) in 78 F1 (SJL/J × B10.S) × B10.S mice. The *Ilg* and *Mdm1* markers were described in Bureau et al. (5). A microsatellite was found in the second intron of *Myf6* and used to design two PCR primers (forward 5'-GAAAGGGCACTGGCTGTAC-3' and reverse 5'-GCGCGGATTGGCTGTGCT-3') (13). Black squares represent homozygous mice; white squares represent heterozygous mice. Numbers under each column represent the number of mice with each genotype.

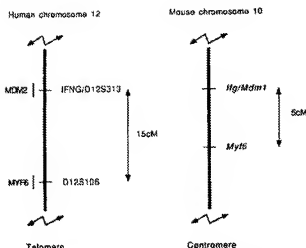


FIG. 3. Genetic maps of the syntenic regions of human chromosome 12q15 and mouse chromosome 10. The MDM2 and MYF6 genes were localized by physical mapping only. The other human genes and markers were localized by both physical and genetic mapping.

of mouse chromosome 10. This information will be extremely useful in human genetic studies in which the IFNG and MDM2 genes will be candidate genes.

#### ACKNOWLEDGMENTS

We thank C. Dib and J. Weissenbach for their help; C. Petit for helpful discussion; and M. Gau for preparing the manuscript. This work was supported by grants from the Institut Pasteur Foundation, the ARSEP, the CNRS, the GREG, and the NMSR.

#### REFERENCES

- Albertsen, H. M., Abderrahim, H., Cann, H. M., Dausset, J., Le Paslier, D., and Cohen, D. (1990). Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. *Proc. Natl. Acad. Sci. USA* 87: 4256-4260.
- Ashar, H. R., Benson, K. F., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., and Chada, K. K. (1994). Hgf, Gfi, Mdm1, Mdm2, and Mdm3: Candidate genes for the mouse pg locus. *Mamm. Genome* 5: 605-611.
- Awata, T., Matsumoto, C., Urakami, T., Hagura, R., Amemiya, S., and Kanazawa, Y. (1994). Association of polymorphism in the interferon gamma gene with IDDM. *Diabetologia* 37: 1159-1162.
- Braun, T., Bober, E., Winter, B., Rosenzthal, N., and Arnold, H. H. (1990). Myf-6, a new member of the human gene family of myogenic determination factors. Evidence for a gene cluster on chromosome 12. *EMBO J.* 9: 821-831.
- Bureau, J. F., Montupetelli, X., Bihl, F., Lefebvre, S., Guenet, J. L., and Brachet, M. (1993). Mapping loci influencing the persistence of Theiler's virus in the murine central nervous system. *Nature Genet.* 5: 87-91.
- Cohen, D., Chumakov, I., and Weissenbach, J. (1993). A first-generation physical map of the human genome. *Nature* 366: 699-701.
- Dietrich, W. F., Miller, J. C., Steen, R. G., Merchant, M., Darron, D., Nahf, R., Gross, A., Joyce, D. C., Wessel, M., Dredge, R. D., Marquis, A., Stein, L. D., Goodman, N., Page, D. C., and Lander, E. S. (1994). A genetic map of the mouse with 4006 simple sequence length polymorphisms. *Nature Genet.* 7: 220-225.
- Fiette, L., Aubert, C., Müller, U., Huang, S., Agut, M., and Brachet, M. (1995). Theiler's virus infection of 129/Sv mice that lack the interferon alpha/beta or interferon gamma receptors. *J. Exp. Med.* 181.
- Francke, U. (1994). Digitized and differentially shaded human chromosome ideograms for genomic applications. *Cytogenet. Cell Genet.* 65: 208-218.
- Guyap, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., and Weissenbach, J. (1994). The 1993-94 Généthon human genetic linkage map. *Nature Genet.* 7: 246-339.
- Highway, J., Mitchell, E. L. D., Jones, D., White, G. R. M., and Santibenz Koref, M. F. (1994). A transcribed polymorphism and sub-localization of MDM2. *Hum. Genet.* 93: 611-612.
- Isaacs, A., Lindenmann, J., and Valentine, R. C. (1957). Virus interference. II. Some properties of interferon. *Proc. R. Soc. London B* 147: 268-273.
- Miner, J. H., and Wold, B. (1990). Herculin, a fourth member of the myoD family of muscle regulatory genes. *Proc. Natl. Acad. Sci. USA* 87: 1089-1093.
- Naylor, S. L., Sakaguchi, A. Y., Shows, T. B., Law, M. L., Goodall, D. V., and Gray, P. W. (1983). Human immune interferon gene is located on chromosome 12. *J. Exp. Med.* 157: 1020-1027.
- Oliner, J. D., Kinzler, K. W., Meltzer, P. S., George, D. L., and Vogelstein, B. (1992). Amplification of a gene encoding a p53-associated protein in human sarcomas. *Nature* 358: 80-83.
- Reifenberger, G., Lui, L., Ichimura, K., Schmidt, E. E., and Collins, V. P. (1993). Amplification and overexpression of the MDM2 gene in a subset of human malignant gliomas without p53 mutations. *Cancer Res.* 53: 2736-2739.
- Ruiz-Linares, A. (1993). Dinucleotide repeat polymorphism in the interferon-gamma (IFNG) gene. *Hum. Mol. Genet.* 2: 1508.
- Taylor, B. A., Rowe, L., and Grieco, D. (1992). Close linkage of Mdm-1, a gene amplified and overexpressed in a transformed 3T3 cell line, with gamma interferon (IFG) on Chromosome 10 of the mouse. *Mamm. Genome* 3: 700-704.
- Trent, J. M., Olson, S., and Lawn, R. M. (1982). Chromosomal localization of human leukocyte, fibroblast, and immune interferon genes by means of *in situ* hybridization. *Proc. Natl. Acad. Sci. USA* 79: 7809-7813.



**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Re Application of: Ryan

Serial No.: 10/608,463

Group Art Unit: 1652

Filed: June 27, 2003

Examiner: E. Slobodyansky

FOR: ISOLATED GENOMIC POLYNUCLEOTIDE FRAGMENTS FROM  
CHROMOSOME 12 THAT ENCODE HUMAN CARBOXYPEPTIDASE M AND THE  
HUMAN MOUSE DOUBLE MINUTE 2 HOMOLOG

Confirmation No.: 6428

**INFORMATION DISCLOSURE STATEMENT**

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

In accordance with 37 C.F.R. 1.56, 1.97 and 1.98, Applicants submit herewith references which they believe may be material to the patentability of this application and with respect to which there may be a duty to disclose in accordance with 37 C.F.R. 1.56.

While the references may be "material" under 37 C.F.R. 1.56, it is not intended to constitute an admission that the references are "prior art" unless specifically designated as such.

The filing of this Information Disclosure Statement shall not be construed as a representation that no other material references than those listed exist or that a search has been conducted.

The references are listed in PTO form 1449 which is in accordance with the requirements of M.P.E.P. 609. A copy of the references is also enclosed.

The references are as follows:

**U.S. Patent Documents**

## Foreign Patent Documents

### Other Documents

1. RIES et al., 2000, Cell 103:321-330
2. REHLI et al., 1995, J. Biol. Chem. 270:15644-9
3. OLINER et al., 1992, Nature 358: 80-3
4. TAN et al., 1989, J. Biol. Chem. 264: 13165-70

It is respectfully requested that these references be considered by the Patent and Trademark Office in its examination of the above-identified application and be made of record therein. The Examiner is also invited to contact the Undersigned if there are any questions concerning this paper or the attached references.

The Information Disclosure Statement submitted herewith is being filed

☐ before the mailing date of a first Office Action on the merits

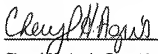
☒ after the mailing date of a first Office Action on the merits. Please charge the fee of \$180 to credit card. Form PTO-2038 is attached.

☐ An International Search Report is enclosed. References cited in the International Search report are asterisked.

Respectfully submitted,

Date:

3/1/05



Cheryl M. Agnis, Reg. No. 34,086  
P.O. Box 806  
Pelham, N.Y. 10803  
(914) 712-0093

## (Use as many sheets as necessary)

Signal 1 at 2

PTO/SB108A (04-03)

Approved for use through 04/30/2003. OMB 0551-0031  
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

### Complete if Known

Application Number	10/608,463
Filing Date	June 27, 2003
First Named Inventor	Ryan
Art Unit	1652
Examiner Name	E. Slobodyansky
Attorney Docket Number	JR10003

## U. S. PATENT DOCUMENTS

[illegible]

## FOREIGN PATENT DOCUMENTS

[illegible]

Examiner  
Signature

Date	Considered
11/1/2011	11/1/2011
11/2/2011	11/2/2011
11/3/2011	11/3/2011
11/4/2011	11/4/2011
11/5/2011	11/5/2011
11/6/2011	11/6/2011
11/7/2011	11/7/2011
11/8/2011	11/8/2011
11/9/2011	11/9/2011
11/10/2011	11/10/2011
11/11/2011	11/11/2011
11/12/2011	11/12/2011
11/13/2011	11/13/2011
11/14/2011	11/14/2011
11/15/2011	11/15/2011
11/16/2011	11/16/2011
11/17/2011	11/17/2011
11/18/2011	11/18/2011
11/19/2011	11/19/2011
11/20/2011	11/20/2011
11/21/2011	11/21/2011
11/22/2011	11/22/2011
11/23/2011	11/23/2011
11/24/2011	11/24/2011
11/25/2011	11/25/2011
11/26/2011	11/26/2011
11/27/2011	11/27/2011
11/28/2011	11/28/2011
11/29/2011	11/29/2011
11/30/2011	11/30/2011
12/1/2011	12/1/2011
12/2/2011	12/2/2011
12/3/2011	12/3/2011
12/4/2011	12/4/2011
12/5/2011	12/5/2011
12/6/2011	12/6/2011
12/7/2011	12/7/2011
12/8/2011	12/8/2011
12/9/2011	12/9/2011
12/10/2011	12/10/2011
12/11/2011	12/11/2011
12/12/2011	12/12/2011
12/13/2011	12/13/2011
12/14/2011	12/14/2011
12/15/2011	12/15/2011
12/16/2011	12/16/2011
12/17/2011	12/17/2011
12/18/2011	12/18/2011
12/19/2011	12/19/2011
12/20/2011	12/20/2011
12/21/2011	12/21/2011
12/22/2011	12/22/2011
12/23/2011	12/23/2011
12/24/2011	12/24/2011
12/25/2011	12/25/2011
12/26/2011	12/26/2011
12/27/2011	12/27/2011
12/28/2011	12/28/2011
12/29/2011	12/29/2011
12/30/2011	12/30/2011
12/31/2011	12/31/2011

EXAMINER: Initial if reference contains, whether or not citation is in conformance with MPPEB 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. <sup>1</sup> Applicant's unique citation designation number (optional). <sup>2</sup> See Kinds Codes of USPTO Patent Documents at [www.uspto.gov](http://www.uspto.gov) or MPPEB 801.04. <sup>3</sup> Error Code that issued the document, by the two-letter code (WIPO Standard 57.3). <sup>4</sup> For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. <sup>5</sup> Kind of document by the appropriate symbols as indicated on the document under WIPO Standard 57.16 if possible. <sup>6</sup> Applicant to place a check mark here if English language Translation is attached.

This collection of information is required by 37 CFR 1.97 and 1.98. This information is required to obtain or retain a benefit by the public which is in the field (and by the USPTO to process) an application. Confidentiality is governed by 35 USC 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, Washington, DC 20231.

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.



PTO/SB/068 (04-03)

Approved for use through 04/30/2003. OMB 0661-0031  
U.S. Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for the TRADEMARK STATEMENT

# **INFORMATION DISCLOSURE STATEMENT BY APPLICANT**

(Use as many sheets as necessary)

## **Complete if Known**

Application Number	10608,463
Filing Date	June 27, 2003
First Named Inventor	Ryan
Art Unit	1652
Examiner Name	E. Stobodyansky
Attorney Docket Number	JR10003

Sheet 2 of 2

## **NON PATENT LITERATURE DOCUMENTS**

Examiner Initials*	Cite No. <sup>1</sup>	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T <sup>2</sup>
		RIES et al., 2000, Cell 103: 321-330.	
		REHLI et al., 1995, J. Biol. Chem. 270: 15644-15649.	
		OLINER et al., 1992, Nature 358:80-83.	
		TAN et al., 1989, J. Biol. Chem. 264: 13165-13170.	

Examiner  
Signature

Date  
Considered

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with need consultation to applicant.  
1 Applicant's unique citation designation number (optional). 2 Applicant is to place a check mark here if English language Translation is attached.  
This collection of information is required by 37 CFR 1.98. This information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 120 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, Washington, DC 20231.

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

## Amplification of a gene encoding a p53-associated protein in human sarcomas

J. D. Oliner\*, K. W. Kinzler†, P. S. Meltzer‡, D. L. George§ & B. Vogelstein\*

\* The Johns Hopkins Oncology Center, 424 North Bond Street, Baltimore, Maryland 21231, USA

† Departments of Pediatrics and Radiation Oncology, University of Michigan Cancer Center, MSRB 11 C5608, 1150 West Medical Center Drive, Ann Arbor, Michigan 48109, USA  
§ Department of Human Genetics, University of Pennsylvania, Clinical Research Building, Philadelphia, Pennsylvania 19104, USA

DESPITE extensive data linking mutations in the p53 gene to human tumorigenesis<sup>1</sup>, little is known about the cellular regulators and mediators of p53 function. MDM2 is a strong candidate for one such cellular protein; the MDM2 gene was originally identified by virtue of its amplification in a spontaneously transformed derivative of mouse BALB/c cells<sup>2</sup> and the MDM2 protein subsequently shown to bind to p53 in rat cells transfected with p53 genes<sup>3,4</sup>. To determine whether MDM2 plays a role in human cancer, we have cloned the human MDM2 gene. Here we show that recombinant-derived human MDM2 protein binds human p53 *in vitro*, and we use MDM2 clones to localize the human MDM2 gene to chromosome 12q13–14. Because this chromosomal position appears to be altered in many sarcomas<sup>5–7</sup>, we looked for changes in human MDM2 in such cancers. The gene was amplified in over a third of 47 sarcomas, including common bone and soft tissue forms. These results are consistent with the hypothesis that MDM2 binds to p53, and that amplification of MDM2 in sarcomas leads to escape from p53-regulated growth control. This mechanism of tumorigenesis parallels that for virally-induced tumours<sup>8,9</sup>, in which viral oncogene products bind to and functionally inactivate p53.

To obtain human complementary DNA clones, a murine MDM2 cDNA probe was used to initiate cDNA walking in a human library (see legend to Fig. 1). Sequence analysis of 25 clones revealed several cDNA forms indicative of alternative splicing. The predominant human form is compared with its murine counterpart in Fig. 1. There was an open reading frame extending from the 5' end of the human cDNA sequence to nucleotide 1,784. Although this signal for translation initiation could not be unambiguously defined, the ATG at nucleotide 312 was considered the most likely position for several reasons. First, the sequence similarity between human and mouse MDM2 declined dramatically upstream of nucleotide 312. Second, an inverse polymerase chain reaction (PCR) was used in an attempt to acquire additional upstream cDNA sequence<sup>10</sup>. The 5' ends of the PCR-derived clones were very similar (within 12 base pairs) to the 5' ends of clones obtained from the cDNA library, indicating that the 5' end of the human MDM2 sequence shown in Fig. 1 may represent the 5' end of the transcript. Third, *in vitro* translation of the sequence shown in Fig. 1, beginning with the methionine encoded by the ATG at position 312, generated a protein similar in size to that observed in human cells (see below).

Comparison of the human and mouse MDM2 coding regions showed that they were 80.3% identical and shared a basic nuclear localization signal at codons 181 to 185 (ref. 11), several casein kinase II serine-phosphorylation sites<sup>12</sup>, an acidic activation domain at codons 223 to 274 (ref. 13), and two metal-binding sites at codons 305 to 322 and 461 to 478, neither of which is highly related to known DNA-binding domains<sup>14</sup>.

To determine whether the human MDM2 protein could bind to human p53 protein *in vitro*, a human MDM2 expression vector was constructed from the cDNA clones (see legend to Fig. 2). RNA transcribed from this vector using T7 RNA polymerase was used to program a rabbit reticulocyte lysate. Although the predicted size of the protein generated from the construct was only 55.2K (*M<sub>r</sub>* 55,200, extending from the methionine at nucleotide 312 to nucleotide 1,784), protein translated *in vitro* migrated at ~90 K. The MDM2 protein was not immunoprecipitated with antibodies against either the C-terminal or N-terminal regions of p53 (Fig. 2, lanes 2 and 3). But when *in vitro*-translated human p53 was mixed with the human MDM2 translation product, the anti-p53 antibodies precipitated MDM2 protein with p53 (Fig. 2, lanes 5 and 6). As a control, a protein of similar electrophoretic mobility (MCC<sup>15</sup>) was mixed with p53 and there was no coprecipitation (Fig. 2, lanes 8 and 9). When an *in vitro*-translated His-175 mutant form of p53 was mixed with human MDM2 protein, a similar coprecipitation of MDM2 and p53 proteins was also observed (data not shown).

Polyclonal rabbit antibodies were raised against an *Escherichia coli*-produced human MDM2-glutathione S-transferase fusion protein. The anti-MDM2 antibodies immunoprecipitated p53 when mixed with MDM2 protein (Fig. 2, lane 15) but failed to precipitate p53 alone (Fig. 2, lane 13).

To establish the chromosomal localization of human MDM2, somatic cell hybrids were screened, and a human-hamster hybrid containing only human chromosome 12 hybridized to the human MDM2 probe. Screening of hybrids containing portions of chromosome 12 (ref. 16) with the same probe narrowed the localization to chromosome 12q13–14. Because this region of chromosome 12 is often aberrant in human sarcomas<sup>5–7</sup>, southern blot analysis to evaluate whether MDM2 was genetically altered in such cancers. We found a striking amplification of MDM2 sequences in several of these tumours (see examples in Fig. 3, lanes 2, 3 and 5). Of 47 sarcomas analysed, 17 showed a 5–50-fold MDM2 amplification. These tumours included 7 of 13 liposarcomas, 7 of 22 malignant fibrous histiocytomas, 3 of 11 osteosarcomas, and 0 of 1 rhabdomyosarcoma. Five benign soft tissue tumours (lipomas) and seventy-four carcinomas (colorectal or gastric) were also analysed by Southern blotting and no amplification was seen.

We next determined whether this gene amplification was associated with increased expression. Because of RNA degradation in primary sarcomas, only the cell lines could be productively analysed by northern blotting. In the one available sarcoma cell line with MDM2 amplification, a single transcript of ~5.5 kilobases (kb) was observed (Fig. 4a, lane 1). The amount of this transcript was much higher than in a sarcoma cell line without amplification (Fig. 4a, lane 2) or in a carcinoma cell line (Fig. 4a, lane 3). When purified messenger RNA (rather than total RNA) from the carcinoma cell line was used for analysis, a human MDM2 transcript of 5.5 kb could also be observed (Fig. 4a, lane 4). Expression of the MDM2 RNA in the sarcoma with amplification was estimated to be at least 30-fold higher than that in the other lines examined. This was consistent with results from western blot analysis. A protein of *M<sub>r</sub>* ~90K was expressed at high levels in the sarcoma cell line with MDM2 amplification (Fig. 4b, lane 3), whereas no expression was evident in two sarcoma cell lines without amplification or in the carcinoma cell line (Fig. 4b, lanes 1, 2 and 4). Five primary sarcomas were also analysed by western blotting. Three primary sarcomas with amplification expressed the same sized protein as that in the sarcoma cell line (Fig. 4c, lanes 1–3), but no protein was observed in the two sarcomas without amplification (Fig. 4c, lanes 4 and 5).

Our results demonstrate that human MDM2 binds to p53 *in vitro* and is genetically altered in a significant fraction of the most common sarcomas of soft tissue and bone<sup>15,16</sup>. It is important to note, however, that amplifications in human tumours

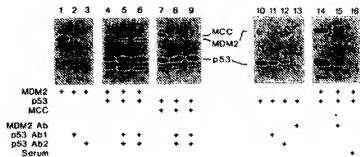
† To whom correspondence should be addressed.





FIG. 2 Coprecipitation of human MDM2 and p53. *In vitro*-translated MDM2, p53 and MCC proteins were mixed as indicated and incubated with p53 Ab1 (monoclonal antibody specific for the C terminus of p53), p53 Ab2 (monoclonal antibody specific for the N terminus of p53), MDM2 Ab (polyclonal rabbit anti-human MDM2 antibodies) or serum (preimmune serum obtained from the rabbit that produced the MDM2 antibody). Lanes 1, 4, 7, 10 and 14 contain aliquots of the protein mixtures used for immunoprecipitation. Bands running slightly faster than p53 are polypeptides produced from internal translation initiation sites.

**Materials and Methods.** A human MDM2 expression vector was constructed in pBluescript SK+ (Stratagene) from overlapping cDNA clones. The construct contained the sequence shown in Fig. 1 from nt 312 to 2176. A 42-bp black beetle virus ribosome entry sequence<sup>25</sup> was placed immediately upstream of this MDM2 sequence in order to obtain high expression. This construct, as well as p53 (ref. 26) and MCC<sup>27</sup> constructs in pBluescript SK+ were transcribed with T7 RNA polymerase and translated in a rabbit reticulocyte lysate (Promega) according to the manufacturer's instructions. Lysate (10 µl) containing the three proteins, alone or mixed in pairs, was incubated at 37 °C for 15 min. 1 µg (10 µl) of p53 Ab1 or Ab2 (Oncogene Science) or 5 µl of rabbit serum containing MDM2 antibody or preimmune rabbit serum were added as indicated. 90 µl RIPA buffer (3.0 mM Tris, pH 7.5, 1% sodium dodecylsulfate, 1% NP40, 150 mM NaCl, 0.1% SDS, SNNT buffer<sup>28</sup> or binding buffer<sup>29</sup>) were then added and the mixtures allowed to incubate at 4 °C for 2 h. The three buffers produced similar results, although the



coprecipitation was less efficient in SNNT buffer (containing 0.6 M NaCl; lanes 5 and 8) than in binding buffer (containing 0.1 M NaCl; lanes 6 and 9). Following addition of 2 mg protein A-Sepharose, the tubes were rotated and over- and at 4 °C for 1 h. After pelleting and washing, immunoprecipitates were electrophoresed on SDS-polyacrylamide gels and the dried gels autoradiographed in the presence of Enhance (New England Nuclear). Rabbits were immunized with a glutathione S-transferase (Pharmacia)-MDM2 fusion protein containing human MDM2 from the region corresponding to nt 390-816.

FIG. 3 Amplification of the human MDM2 gene in sarcomas. DNA (5 µg) was digested with *EcoRI*, separated by agarose gel electrophoresis and transferred to nylon as described<sup>27</sup>. Filters were then hybridized with a human MDM2 cDNA fragment probe (nt 1-949; Fig. 1) or to a control probe that identifies fragments of similar size (pOCG 1.65; ref. 28). Hybridization was as previously described<sup>28</sup>. DNA was derived from 5 primary sarcomas (lanes 1-4, 6) and one sarcoma cell line (Osa-CL; lane 5). On longer exposure, the same sized MDM2 fragments were observed in lanes 1, 4 and 6. DNA fragment sizes are shown on the left in kb.

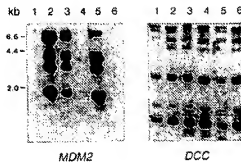
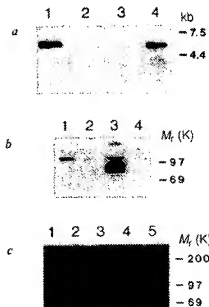


FIG. 4 MDM2 expression. a Northern blot analysis. RNA was separated by electrophoresis in a MOPS-formaldehyde gel and electrophoretically transferred to nylon filters. Transfer and hybridization were as described<sup>26</sup>. RNA was hybridized to the MDM2 fragment described in Fig. 3 legend. Total RNA (10 µg) was derived, respectively, from two sarcoma cell lines (Osa-CL; lane 1) and RC13; lane 2) and the colorectal cancer cell line (CaCo-2) used to make the cDNA library (lane 3). Lane 4 contains 10 µg polyadenylated CaCo-2 RNA. RNA sizes are shown on the right in kb. b Western blot analysis of the sarcoma cell lines RC13 (lane 1), Osa-CL (lane 3), HOS (lane 4), and the carcinoma cell line CaCo-2 (lane 2). c Western blot analysis of primary sarcomas. Lanes 1 to 3 contain protein from sarcomas with MDM2 amplification, and lanes 4 and 5 contain protein from sarcomas without MDM2 amplification. Western blots using affinity-purified MDM2 antibody were performed with 50 µg protein per lane as described<sup>30</sup>, except that the membranes were blocked in 10% non-fat dried milk and 10% goat serum, and secondary antibodies were coupled to horseradish peroxidase to allow chemiluminescent detection (Amersham ECL). MDM2 antibody was affinity-purified with a pATH-MDM2 fusion protein using methods described in ref. 31. Nonspecific reactive proteins of 75, 105 and 170K are seen in all lanes, irrespective of MDM2 amplification. MDM2 proteins, of M<sub>r</sub> 90K, were observed only in the MDM2-amplified tumours. Protein marker sizes are shown on the right.



BEST AVAILABLE COPY

(unpublished results with T. Tokino and D. Sidransky). The amplification of *MDM2* provides another provocative parallel between viral carcinogenesis and the naturally occurring genetic alterations underlying sporadic human cancer. □

Received 20 March; accepted 5 May 1992.

1. Holstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. *Science* **253**, 499–501 (1991).
2. Yip-Schindler, S., Tsai, S.-Y., Chen, C. & George, D. *Oncogene* **5**, 1545–1549 (1991).
3. Hanks, P. W. *et al.* *Cell Growth Differ.* **3**, 571–580 (1992).
4. Mondorf, J., Zambetti, G. P., Olson, D. C., George, D. L. & Levine, A. J. *Cell* (in the press).
5. Mendon, N. *et al.* *Cancer Chemother. Oncol.* **4**, 7–14 (1989).
6. Tur-Castell, C. *et al.* *Cancer gene. Cytogenet.* **33**, 291–299 (1985).
7. Melzer, R. P. *et al.* *Cell Growth Differ.* **2**, 495–501 (1991).
8. Lane, D. P. & Baltimore, D. *Science* **250**, 4–8 (1990).
9. Werner, B. A., Levine, A. J. & Hawley, P. M. *Science* **248**, 70–73 (1990).
10. Ochman, H., Alikhan, J., Garcia, D. & Hartl, D. *in* *PCR Technology: Principles and Applications for DNA Amplification* (eds White, J. & Maniatis, T.) (Cold Spring Harbor, New York, 1990).
11. Tanaka, K. *et al.* *EMBO J.* **10**, 41–45 (1991).
12. Pines, L. A. *Biochem. Biophys. Res. Commun.* **180**, 267–274 (1991).
13. Pines, L. A. *Nature* **353**, 553–556 (1991).
14. Herrmann, S. C. *Nature* **353**, 715–719 (1991).
15. Kunkel, K. W. *et al.* *Science* **251**, 1368–1370 (1991).
16. Low, M. L. *et al.* *Ann. Hum. Genet.* **56**, 121–137 (1992).
17. Winkler, S. W. & Doolittle, R. F. *Mol. Cancer* **4**, 2250–2266 (1978).
18. Mowles, M.A., Abelson, H. T. & Sait, M. D. *in* *Cancer: Principles and Practice of Oncology* (eds DeVita, T. L., Hellman, S. & Rosenberg, S. A.) (1992) 1342, Lippincott, Philadelphia, 1993).
19. Hanks, P. W. *et al.* *Science* **256**, 70–73 (1991).
20. Kunkel, K. W. *et al.* *Proc. Natn. Acad. Sci. USA* **88**, 1031–1035 (1991).
21. Brown, D. H. & Sanger, H. C. *Cancer gene. Cytogenet.* **35**, 104–121 (1986).
22. Guber, U. & Hoffmann, B. *J. Gene* **25**, 203–206 (1983).
23. Ellegren, J., Mulligan, J. T., Pomeroy, S. W., Spentwood, M. & Davis, R. W. *Proc. Natn. Acad. Sci. USA* **88**, 1771–1785 (1991).
24. Feenberg, A. & Vogelstein, B. *Analyt. Biochem.* **182**, 81–83 (1989).
25. Damschroder, B., Rothen, E. J. & Schwartz, J. *Nucleic Acids Res.* **16**, 3633 (1987).
26. Dandry, W. S., Korn, S. E., Peterson, J. A., Kunkel, K. W. & Vogelstein, B. *Nature Genet.* **1**, 45–49 (1990).
27. Reed, K. C. & Mann, D. A. *Nucleic Acids Res.* **18**, 7207–7215 (1990).
28. Taylor, E. R. *et al.* *Science* **247**, 45–48 (1989).
29. Vogelstein, B. *et al.* *Cancer Res.* **47**, 4018–4021 (1987).
30. Kunkel, K. W., Ruppert, J. M., Rigler, S. H. & Vogelstein, B. *Nature* **332**, 371–374 (1988).
31. Kunkel, K. W. & Vogelstein, B. *Molec. cell. Biol.* **10**, 634–642 (1990).

**ACKNOWLEDGMENTS.** We thank A. J. Levine for the information about *MDM2*-p53 interaction that stimulated this work. M. L. Lane for somatic cell hybrids. S. Ellegren for the lambda *SV40* vector, and T. Oshiro for preparation of the manuscript. This work was supported by the Preuss Foundation, the Clayton Fund, and grants from the National Institutes of Health.

## Wild-type p53 activates transcription *in vitro*

George Farmer, Jill Bargonetti, Hua Zhu, Paula Friedman, Ron Prywes & Carol Prives

Department of Biological Sciences, Columbia University, New York 10027, USA

THE p53 protein is an important determinant in human cancer and regulates the growth of cells in culture<sup>1–3</sup>. It is known to be a sequence-specific DNA-binding protein<sup>4,5</sup> with a powerful activation domain<sup>6–8</sup>, but it has not been established whether it regulates transcription directly. Here we show that intact purified wild-type human and murine p53 proteins strongly activate transcription *in vitro*. This activation depends on the ability of p53 to bind to a template bearing a p53-binding sequence. By contrast, tumour-derived mutant p53 proteins cannot activate transcription from the template at all, and when complexed to wild-type p53, these mutants block transcriptional activation by the wild-type protein. Moreover, the simian virus 40 large T antigen inhibits wild-type p53 from activating transcription. Our results support a model in which p53 directly activates transcription, but this activity can be inhibited by mutant p53 and SV40 large T antigen through interaction with wild-type p53.

A DNA-binding-immunoassay has been used to screen human genomic clones and show that p53 binds specifically to a region upstream of the transcription start site for the human ribosomal gene cluster (RGC)<sup>9</sup>. We have confirmed and extended this observation by DNase I footprinting and shown that addition of immunopurified p53 to a DNA fragment containing the RGC

site leads to strong and specific protection of only the RGC region (Fig. 1). All tumour-derived mutant p53 proteins tested failed to protect this sequence (J.B. *et al.*, manuscript in preparation).

To determine whether p53 can activate transcription *in vitro*, we used as templates the plasmids *foswt* and *foswt*, which contain the human RGC p53 DNA-binding fragment or a mutated RGC fragment respectively (Fig. 2a). Three partially purified fractions from HeLa cell nuclear extracts were used as a source of transcription factors<sup>9</sup>. RNA products were analysed by S1 nuclease digestion using specific probes for each construct. Increasing amounts of p53 stimulated transcription from *foswt* but (compare lanes 1–4 with lanes 5–8). These reaction mixtures also included a construct containing an abridged adenovirus major late promoter (pML5; ref. 10) whose transcription was not significantly affected by p53.

The p53 protein activated transcription from another promoter as well (Fig. 2b). Plasmids containing either one or sixteen copies of the RGC site, or one mutant RGC site, inserted adjacent to the polyoma virus early promoter to create Pylwt, Pyl6wt and Pylmt, respectively, were used as templates in transcription reactions. We found that p53 activated transcription of constructs containing the wild-type RGC (lanes 1–9) but not the mutant RGC (lanes 10–12). Diagrams of the templates and the test probe used in these experiments are shown in Fig. 2c with the expected S1 nuclease products.

The high incidence of p53 gene mutations in cancer patients suggests that alteration of the normal function of p53 is an important part of the oncogenic process. Therefore it was of interest to examine whether tumour-derived mutant p53 proteins activate transcription. The mutant p53 proteins we chose are defective in both nonspecific and specific DNA binding<sup>9,11</sup>, so providing an opportunity to confirm that p53 must bind DNA to activate transcription. We compared the ability of wild-type and two tumour-derived mutant p53 proteins with mutations at either amino acid 175 (His 175) or at amino acid 273 (His 273)



## Molecular Cloning and Sequencing of the cDNA for Human Membrane-bound Carboxypeptidase M

COMPARISON WITH CARBOXYPEPTIDASES A, B, H, AND N\*

(Received for publication, April 21, 1989)

Fulong Tan†, Shu Jin Chan‡, Donald F. Steiner§, James W. Schilling¶, and Randal A. Skidgel‡

From the Laboratory of Peptide Research and Departments of Pharmacology and Anesthesiology, University of Illinois College of Medicine, Chicago, Illinois 60608, the §Howard Hughes Medical Institute and Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, and the ¶California Biotechnology, Inc., Mountain View, California 94043

Carboxypeptidase M, a widely distributed membrane-bound carboxypeptidase that can regulate peptide hormone activity, was purified to homogeneity from human placenta (Skidgel, R. A., Davis, R. M., and Tan, F. (1989) *J. Biol. Chem.* 264, 2236-2241). The NH<sub>2</sub>-terminal 31 amino acids were sequenced, and two complementary oligonucleotide probes were synthesized and used to isolate a carboxypeptidase M clone from a human placental cDNA library. Sequencing of the cDNA insert (2009 base pairs) revealed an open reading frame of 1317 base pairs coding for a protein of 439 residues. The NH<sub>2</sub>-terminal protein sequence matched the deduced amino acid sequence starting with residue 14. Hydrophobic analysis revealed hydrophobic regions at the NH<sub>2</sub> and COOH termini. The NH<sub>2</sub>-terminal 13 amino acids probably represent part of the signal peptide, and the COOH-terminal hydrophobic region may act either as a transmembrane anchor or as a signal for attachment to a phosphatidylinositol glycan moiety. The carboxypeptidase M sequence contains six potential Asn-linked glycosylation sites, consistent with its glycoprotein nature. The sequence of carboxypeptidase M was 41% identical with that of the active subunit of human plasma carboxypeptidase N, 41% identical with bovine carboxypeptidase H (carboxypeptidase E, enkephalin convertase), and 15% with either bovine pancreatic carboxypeptidase A or B. Many of the active site residues identified in carboxypeptidase A and B, including all of the zinc-binding residues (2 histidines and a glutamic acid), are conserved in carboxypeptidase M. These data indicate that all of the metallo-carboxypeptidases are related, but the nondigestive carboxypeptidases with more specialized functions, present in cell membranes, blood plasma, or secretory granules (i.e., carboxypeptidase M, carboxypeptidase N and carboxypeptidase H), are more closely related to each other (41-49% identity) than they are to carboxypeptidase A or B (15-20% identity).

### Carboxypeptidases which specifically remove COOH-ter-

\*This study was supported by Grants DK13914 (to D. F. S.), HL36473, and HL36082 from the National Institutes of Health and by the Howard Hughes Medical Institute. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact. The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J04970.

minal basic amino acids (Arg or Lys) have important functions in many biological processes, including protein digestion, activation, inactivation or modulation of peptide hormone activity, and alteration of the physical properties of proteins and enzymes (1). The actual role that the various mammalian arginine/lysine carboxypeptidases play *in vivo* is probably related to their localization as well as their physical properties. For example, pancreatic carboxypeptidase B (EC 3.4.17.2) is not normally found outside the pancreas or small intestine, except in cases of acute pancreatitis when low levels can be detected in blood (2). The half-life of carboxypeptidase B injected into the bloodstream is on the order of minutes (3) because of its low molecular weight and lack of glycosylation, consistent with its major function in protein and peptide degradation in the digestive tract. In contrast, human plasma carboxypeptidase N (kininase I, anaphylatoxin inactivator, arginine carboxypeptidase, EC 3.4.17.3) circulates in plasma as a large ( $M_r = 280,000$ ) tetrameric complex of two active subunits ( $M_r = 48,000-55,000$ ) and two glycosylated inactive subunits ( $M_r = 53,000$ ) which stabilize the active subunits and keep them in the circulation (4-6). One of the functions of carboxypeptidase N is to protect the body from potent vasoactive and inflammatory peptides containing COOH-terminal Arg or Lys (e.g. kinins, anaphylatoxins) which are released into the circulation (1, 4, 7).

Carboxypeptidase H (also known as enkephalin convertase or carboxypeptidase E, EC 3.4.17.10) is an arginine/lysine carboxypeptidase with an acid pH optimum and is located in secretory granules of pancreatic islets, adrenal gland, pituitary, and brain (8-12). This enzyme probably removes the residual COOH-terminal Arg or Lys remaining after initial endoprotease cleavage during prohormone processing at the intragranular acid pH (1, 12).

Carboxypeptidase M is a membrane-bound arginine/lysine carboxypeptidase found in many tissues and cultured cells (1, 13-15). It was recently purified to homogeneity from human placenta (16). Because of its presence on plasma membranes and optimal activity at neutral pH, it is ideally situated to act on peptide hormones at local tissue sites where it could control their activity before or after interaction with specific plasma membrane receptors (1). Here we report the cloning and sequencing of the cDNA for human carboxypeptidase M. The results show that carboxypeptidase M is a unique enzyme which exhibits similarity to carboxypeptidases A, B, H, and N, and its sequence is consistent with the properties we determined for the purified protein (16).

### EXPERIMENTAL PROCEDURES

**Materials**—Restriction enzymes were purchased from Boehringer Mannheim or New England Biolabs. [ $\gamma$ -<sup>32</sup>P]ATP and <sup>32</sup>S-ATP were

from Amersham Corp., and nitrocellulose filters were from Schleicher & Schuell. All other solvents and chemicals were of reagent grade or better.

**Purification and NH<sub>2</sub>-terminal Sequence Analysis of Human Carboxypeptidase M.**—Carboxypeptidase M was purified to homogeneity from human placental microvilli as described (16) by sequential column chromatography on DEAE-Trisacryl, Q-Sepharose, sepharose, and Mono-Q HR in high performance liquid chromatography. Samples (50–60 µg) of native carboxypeptidase M and carboxypeptidase M reduced with dithiothreitol and alkylated with vinyl pyridine (17) were dialyzed against 5% acetic acid and subjected to sequencing on an Applied Biosystems sequencer (model 477).

**Oligonucleotide Synthesis and Labeling.**—Oligonucleotides used for screening cDNA libraries and for sequencing were synthesized on an Applied Biosystems model 380B DNA synthesizer. After complete deprotection, oligonucleotide probes used for screening were further purified by electrophoresis and elution from 20% polyacrylamide gels and then 5' end-labeled with [ $\gamma$ -<sup>32</sup>P]ATP using T4 polynucleotide kinase (18).

**Screening of cDNA Libraries.**—Two human placental cDNA libraries, a  $\lambda$ gt10 library obtained from Dr. P. Seeburg of Genetech, and a  $\lambda$ gt11 library purchased from Clontech, Palo Alto, CA, were screened using standard techniques (18) with two labeled synthetic oligonucleotide probes complementary to two adjacent regions of the NH<sub>2</sub>-terminal amino acid sequence. Triplicate filter lifts were made from each plate and hybridized with each probe separately or with a mixture of both probes.

**Subcloning and DNA Sequence Analysis.**—Positive plaques were purified by reextracting at low density, and DNA was prepared by the plate lysis method (18). The DNA was cleaved with EcoRI, the 2-kilobase pair (kb)<sup>1</sup> insert was purified by agarose gel electrophoresis and then subcloned into the EcoRI site of pGEM4Z or M13mp18. Large-scale preparation of plasmid or phage DNA was performed as described (18). HindIII restriction fragments (640, 240, and 1170 base pairs (bp)) were run on agarose gel electrophoresis and subcloned into the EcoRI/HindIII site of M13mp18 (19). Single-stranded DNA was prepared and was sequenced using the dideoxy chain termination method (19) employing <sup>32</sup>S-dATP (1000 Ci/mmol, Du Pont-New England Nuclear) and Sequences (United States Biochemical Corp., Cleveland, OH) according to the manufacturer's instructions. In some regions, specific synthetic oligonucleotides were used to prime the chain termination reaction. The cDNA clone for carboxypeptidase M was completely sequenced on both strands. Nucleotide and amino acid sequences were analyzed with the DNASTAR software package (DNASTAR, Madison, WI).

**Northern Analysis.**—Total RNA was isolated from human placenta, kidney, and liver by the guanidine thiocyanate method (19). The RNA (30 µg) was separated by electrophoresis on a 1.2% agarose gel containing 0.8 M formaldehyde (18), transferred to nitrocellulose, and hybridized using a nick-translated <sup>32</sup>P-labeled mixture of the 240- and 640-bp HindIII restriction fragments of carboxypeptidase M cDNA (19). Human placental poly(A)<sup>+</sup>mRNA was isolated from total RNA by oligo(dT)-cellulose chromatography (19) and 0.1–1 µg used in the Northern analysis as above.

## RESULTS AND DISCUSSION

**NH<sub>2</sub>-terminal Sequencing and Isolation of a Clone Encoding Human Carboxypeptidase M.**—Human carboxypeptidase M was purified from placenta, and the following sequence of the first 31 amino acids was obtained on a gas-phase sequencer: Leu-Asp-Phe-Asn-Tyr-His-Arg-Gln-Glu-Gly-Met-Glu-Ala-Phe-Leu-Lys-Thr-Val-Ala-Gln-Xaa-Tyr-Ser-Ser-Val-Thr-His-Leu-His-Ser-Ile. No amino acid was detected at position 21, probably due to the presence of a glycosylated asparagine residue (see below). Two oligonucleotide probes (17-mers, each with 32-fold degeneracy), complementary to adjacent regions of the NH<sub>2</sub>-terminal sequence were synthesized: 5' GA/C/T/T/T/C/T/AAC/T/T/A/C/T/T/C/A/C/T/C/G 3', corresponding to Asp<sup>1</sup>-Arg<sup>2</sup>, and 5' CA/G/A/G/G/A/GG/A/G/C/T/TATGGA/G/AGC 3', corresponding to Gln<sup>24</sup>-Ala<sup>25</sup>. These probes were labeled with [ $\gamma$ -<sup>32</sup>P]ATP and used to screen approximately 500,000 plaques from the two human placental

cDNA libraries. From the initial screening, two positive clones which hybridized with both probes were obtained: ACPM, and ACPM<sub>2</sub>. Both clones contained a 2-kb insert, gave identical restriction maps with BamHI and HindIII, and were therefore considered to be the same. All further work was carried out using ACPM<sub>2</sub>.

**Sequence Analysis of Human Carboxypeptidase M cDNA.**—The cDNA was sequenced completely in both strands using the strategy outlined in Fig. 1. Sequence analysis revealed the insert is 2009 bp long with an open reading frame of 1517 bp coding for a protein of 439 amino acids (Fig. 2). The cDNA did not contain the initiator Met, but the amino acid sequence predicted by the cDNA matches exactly the partial NH<sub>2</sub>-terminal protein sequence determined for carboxypeptidase M, indicating the clone is authentic and that it codes for the entire mature protein. Hydrophobic analysis of the sequence reveals two hydrophobic regions; one at the NH<sub>2</sub> terminus and one at the COOH terminus (Fig. 2). The NH<sub>2</sub>-terminal 13 amino acids probably represent a portion of the signal peptide, as the NH<sub>2</sub>-terminal sequence of the mature protein does not contain this sequence (Figs. 2 and 4). In addition, as was shown for other signal peptides (20), it is hydrophobic and contains amino acids with small side chains at the -1 and -3 positions relative to the potential signal peptide cleavage site (Fig. 2). Thus, the mature protein would consist of 426 residues with a calculated molecular weight of 48,709, in good agreement with the value of 47,600  $\pm$  1,000 determined for deglycosylated carboxypeptidase M in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (16).

**Potential Membrane Binding Region.**—Carboxypeptidase M is tightly bound to plasma membranes in various tissues and cells (1, 13–15). As a membrane enzyme, it could be anchored by either a hydrophobic transmembrane segment or by covalent attachment to phosphatidylinositol groups (21–23). Because the NH<sub>2</sub>-terminal hydrophobic sequence is a cleaved signal peptide, the COOH-terminal hydrophobic region (residues 411–425) is the most likely domain to function as either a membrane anchor or as a signal for phosphatidylinositol glycan attachment (Figs. 2 and 3). This stretch of 15 residues is flanked by charged lysines (Figs. 2 and 4) as is common with membrane-spanning domains, but it is somewhat shorter than the putative sequence of 20 hydrophobic and uncharged amino acids thought to be necessary to span the membrane as an  $\alpha$  helix (21). However, at least two other membrane-bound proteins, the  $\beta$  subunit of the T-cell antigen receptor (24) and angiotensin I-converting enzyme (25) contain COOH-terminal membrane-spanning regions of only 17 amino acids, and deletion mutation studies of stomatitis virus glycoprotein (26) and bacteriophage  $\phi$  III gene III protein (27) have shown that a stretch of 12–14 amino acids is sufficient

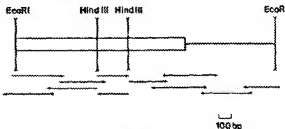


FIG. 1. Restriction map and sequencing strategy of the carboxypeptidase M cDNA clone. The 2.0-kb insert is shown with the coding region represented by the open rectangle. Arrows show the origins and directions of sequencing. Arrows with asterisks indicate sequencing reactions which were primed with specific oligonucleotide primers.

<sup>1</sup> The abbreviations used are: kb, kilobase pair; bp, base pairs.



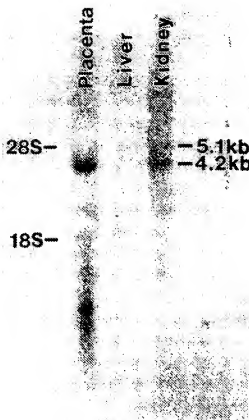


Fig. 5. Northern analysis of total RNA from human tissues. Total RNA (20  $\mu$ g) from human placenta, liver, or kidney was separated on an agarose-formaldehyde gel, blotted onto nitrocellulose, and probed with a  $^{32}$ P-labeled mixture of the 240- and 640-bp *Hind*III restriction fragments of carboxypeptidase M cDNA.

labeled mixture of the 240- and 640-bp *Hind*III restriction fragments of carboxypeptidase M cDNA (Fig. 5). A 24-h exposure revealed a single mRNA species of 4.2 kb in the placenta and two light bands of 4.2 and 5.1 kb in the kidney (Fig. 5). After longer exposure (2 weeks), two light bands were seen in the liver RNA sample which were the same size as those in the kidney (not shown). A repeat of the Northern blot on poly(A)<sup>+</sup> RNA isolated from the placenta gave the same 4.2-kb band seen in the blot of total RNA (not shown). The relative amounts of mRNA are consistent with the levels of carboxypeptidase M activity we have found in these tissues: it is lowest in the liver,<sup>2</sup> higher in kidney, and highest in the placenta (13). However, the size of the mRNA in the placenta is significantly larger than the 2-kb cloned cDNA we have isolated. Although the 3' end of the cDNA clone contains a poly(A) tract, it appears not to represent true poly(A) tailing, as a polyadenylation signal was not found (Fig. 2). Moreover, we have recently isolated a genomic clone containing human carboxypeptidase M, and partial sequence analysis revealed the presence of the same poly(A) tract.<sup>2</sup> Thus, the carboxypeptidase M mRNA likely contains a larger 3'-untranslated region than shown in Fig. 2 and also a 5'-untranslated region of undetermined length. The additional 5.1-kb band in the kidney and liver mRNA could arise from alternative splicing,



Fig. 6. Comparison of the sequences of human carboxypeptidases M and N and bovine carboxypeptidases H, B, and A. For human carboxypeptidase N, the sequence given is for the lower molecular weight active subunit. Residues which are identical in at least three of the five sequences are boxed. Gaps were introduced to optimize the alignment. Amino acids thought to be important in the active centers of bovine carboxypeptidases A and B are marked as follows: solid triangles, zinc-binding residues; solid circles, catalytic residues; solid squares, substrate binding residues. Abbreviations: H, human; CP, carboxypeptidase.

the use of multiple transcription start sites or multiple polyadenylation signals.

**Comparison with Other Carboxypeptidases.**—A computer search of the protein data base revealed significant similarity of the sequence of human carboxypeptidase M to those of bovine carboxypeptidases A (29), B (30), and H (31). There was also a high degree of similarity to the partial protein

<sup>2</sup> R. M. Davis, P. A. Deddish, and R. A. Skidgel, unpublished results.

<sup>3</sup> F. Tan and S. Chan, unpublished results.

TABLE I

Amino acid sequence identity of mammalian carboxypeptidases

The values given are percent identity, calculated by dividing the number of identical residues by the number of amino acids in the longer of the two sequences compared. Abbreviations: H, human; B, bovine; CP, carboxypeptidase. H-CPN refers to the sequence of the active subunit of human carboxypeptidase N.

	H-CPM	H-CPN	B-CPH	B-CPB	B-CPA
			%		
H-CPM		41	41	15	15
H-CPN			49	17	19
B-CPH				17	20
B-CPB					48

sequence of the active subunit of human carboxypeptidase N (32) and to the full sequence (deduced from the cDNA) published recently (33). Fig. 6 shows the alignment of the five carboxypeptidases. The overall identity of human carboxypeptidase M was highest (41%) with bovine carboxypeptidase H and the active subunit of human plasma carboxypeptidase N, whereas the identity with bovine pancreatic carboxypeptidase A or B was much less (Table I). Certain regions are highly conserved (89–93% identity) between carboxypeptidases M, N, and H. These are residues 57–75, 100–114, 126–141, 166–184, and 258–276 (numbered according to carboxypeptidase M; Fig. 6). Most of these regions contain residues which correspond to active site residues identified in bovine carboxypeptidase A and B. Thus, the three zinc-binding residues (His<sup>60</sup>, Glu<sup>72</sup>, and His<sup>104</sup> in carboxypeptidase A), two of the substrate-binding residues (Arg<sup>161</sup> and Tyr<sup>183</sup>), and the catalytic glutamic acid (Glu<sup>276</sup>) are strictly conserved in all of the carboxypeptidases (Fig. 6). Other residues presumed to be involved in catalysis, such as Ser<sup>197</sup> and Tyr<sup>198</sup> of bovine carboxypeptidase A, are not conserved in carboxypeptidases M, N, and H which all contain glycine residues in these positions (Fig. 6). However, it was proposed for a recent model of carboxypeptidase A (34) that the carbonyl oxygen of Ser<sup>197</sup> functions to stabilize the tetrahedral intermediate of the substrate and that the side chains of Ser<sup>197</sup> and Tyr<sup>198</sup> are not involved in catalysis. Thus, the carbonyl oxygen of Gly may be able to fulfill this role in carboxypeptidases M, N, and H. In carboxypeptidase A, Arg<sup>237</sup> is thought to polarize the scissile carbonyl group in the substrate (34). Carboxypeptidases M, N, and H contain Lys in this position (Fig. 6) which may have the same function. Because carboxypeptidases B, M, N, and H all specifically cleave COOH-terminal basic amino acids (Arg or Lys), a negatively charged active site residue would be required to bind the positively charged substrate side chain. In carboxypeptidase B, the substrate side-chain-binding residue is Asp<sup>253</sup>. Although carboxypeptidases M, N, and H have Glu in this position (which could not function in the same way), all three enzymes contain Asp as the next residue, which might serve the same role (Fig. 6).

The regions of sequence identity between carboxypeptidases M, N, and H are spread throughout the proteins with one exception: the COOH-terminal regions have essentially no similarity (Fig. 6). This is probably due to the differences in distribution and functions of the enzymes. As stated above, the COOH-terminal part of carboxypeptidase M probably serves as either a transmembrane anchor or as a signal for attachment to phosphatidylinositol glycan, whereas in carboxypeptidase H the COOH-terminal portion may mediate binding to granule membranes by forming an amphipathic helix (31). In carboxypeptidase N, the COOH-terminal region contains many basic residues which can explain the rapid conversion of the  $M_r = 55,000$  form of the active subunit to an  $M_r = 48,000$  form by various serine proteases (5, 6, 33).

When the sequence identities of all five carboxypeptidases are compared (Table I), they can be divided into two groups: 1) the digestive carboxypeptidases A and B and 2) carboxypeptidases H, M, and N which are present in secretory granules, cell membranes, or blood plasma and thus have different functions. The sequence identity of the enzymes within each group ranges from 41 to 49%, whereas the sequence identity between the two groups is only 15–20% (Table I). Thus, it is possible that an ancestral carboxypeptidase gene duplicated and diverged to evolve into two separate precursor genes, one of which gave rise to carboxypeptidases A and B and the other to carboxypeptidases H, M, and N. A second carboxypeptidase A-type enzyme was recently discovered in rat pancreas which has 63% sequence identity with either rat or bovine carboxypeptidase A, but so far it has not been found in any other species (35). Although carboxypeptidases A and B participate in protein and peptide degradation in the digestive tract, carboxypeptidases H, M, and N have evolved to perform more specialized functions. These include prohormone processing in the acid environment of secretory granules (carboxypeptidase H) or regulation of peptide hormone activity at neutral pH in the blood (carboxypeptidase N) or at the cell surface (carboxypeptidase M) (1, 4, 12).

**Acknowledgments**—We thank Paul Gardner for synthesizing the oligonucleotides used and Dr. Ervin G. Erdős for encouragement and helpful discussions.

## REFERENCES

- Skidgel, R. A. (1989) *Trends Pharmacol. Sci.* **9**, 299–304
- Deik, A. S., Durie, P. R., Fletcher, T. S., and Largman, C. (1965) *Clin. Chem.* **31**, 1294–1300
- Erdős, E. G., Renfrew, A. G., Sliane, E. M., and Wehler, J. R. (1963) *Ann. N. Y. Acad. Sci.* **104**, 222–234
- Erdős, E. G. (ed) (1979) in *Handbook of Experimental Pharmacology*, Vol. 25, Supplement, pp. 428–487, Springer-Verlag, Heidelberg, Federal Republic of Germany
- Levin, Y., Skidgel, R. A., and Erdős, E. G. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4618–4622
- Plummer, T. H., Jr., and Hurwitz, M. Y. (1978) *J. Biol. Chem.* **253**, 3907–3912
- Huey, R., Bloor, C. M., Kawahara, M. S., and Hugli, T. E. (1983) *Am. J. Pathol.* **112**, 48–60
- Zubik, H., Steiner, D. P., Lemmark, A., and Lipsky, C. (1975) *Chrom. Syn. 41*, 183–195
- Davidson, H. W., and Hutton, J. C. (1987) *Biochem. J.* **245**, 575–582
- Hook, V. Y., and Loh, Y. P. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2776–2780
- Fricker, L. D., and Snyder, S. H. (1983) *J. Biol. Chem.* **258**, 10950–10955
- Fricker, L. D. (1988) *Ann. Rev. Physiol.* **50**, 309–321
- Skidgel, R. A., Johnson, A. R., and Erdős, E. G. (1984) *Biochem. Pharmacol.* **33**, 3471–3478
- Johnson, A. R., Skidgel, R. A., Gafford, J. T., and Erdős, E. G. (1984) *Peptides* **5**, 789–796
- Deddish, P., Skidgel, R. A., Becker, R. P., and Erdős, E. G. (1989) *FASEB J.* **3**, A1026
- Skidgel, R. A., Davis, R. M., and Tan, F. (1989) *J. Biol. Chem.* **264**, 2236–2241
- Tarr, G. E. (1986) in *Methods of Protein Microcharacterization* (Shively, J. E., ed) pp. 155–194, Humana Press, Clifton, NJ
- Mauviel, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Davis, L. G., Dittmer, D. D., and Batty, J. F. (1986) *Basic Methods in Molecular Biology*, Elsevier Scientific Publishing Co., Inc., New York
- von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99–105
- Wickner, W. T., and Lodish, H. F. (1985) *Science* **230**, 400–407
- Low, M. G. (1987) *Biochem. J.* **244**, 1–13
- Low, M. G. (1989) *FASEB J.* **3**, 1600–1608
- Yanagi, Y., Yoshikai, Y., Leggett, K., Clark, S. P., Alexander, I.,

- and Mak, T. (1984) *Nature* **308**, 145-149
25. Soubrier, F., Aihenc-Gelas, F., Hubert, C., Allegrini, J., John, M., Tregear, G., and Corvol, P. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 9386-9390
26. Adams, G. A., and Rose, J. K. (1985) *Cell* **41**, 1007-1015
27. Davis, N. G., Boeke, J. D., and Model, P. (1985) *J. Mol. Biol.* **181**, 111-121
28. Hooper, N. M., and Turner, A. J. (1988) *FEBS Lett.* **229**, 340-344
29. Bradshaw, R. A., Ericsson, L. H., Walsh, K. A., and Neurath, H. (1969) *Proc. Natl. Acad. Sci. U. S. A.* **63**, 1369-1394
30. Titani, K., Ericsson, L. H., Walsh, K. A., and Neurath, H. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 1666-1670
31. Fricker, L. D., Evans, C. J., Each, F. S., and Herbert, E. (1986) *Nature* **323**, 461-464
32. Skidgel, R. A., Bennett, C. D., Schilling, J. W., Tan, F., Weerasinghe, D. K., and Erdős, E. G. (1986) *Biochem. Biophys. Res. Commun.* **164**, 1323-1329
33. Gebhard, W., Schube, M., and Eulitz, M. (1989) *Eur. J. Biochem.* **178**, 603-607
34. Christianson, D. W., David, P. R., and Lipscomb, W. N. (1967) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 1512-1515
35. Gerdel, S. J., Craik, C. S., Clauser, E., Goldsmith, E. J., Stewart, C.-B., Graf, M., and Rutter, W. J. (1983) *J. Biol. Chem.* **258**, 17826-17836
36. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105-132



## Carboxypeptidase M Is Identical to the MAX.1 Antigen and Its Expression Is Associated with Monocyte to Macrophage Differentiation\*

(Received for publication, March 17, 1995, and in revised form, May 1, 1995)

Michael Rehli, Stefan W. Krause, Marina Kreutz, and Reinhard Andreessen†

From the Department of Hematology and Oncology, University of Regensburg, 93042 Regensburg, Germany

The two monoclonal antibodies MAX.1 and MAX.11 recognize cell surface antigens that are almost undetectable on monocytes but highly expressed on differentiated macrophages. Biochemical characterization revealed that both antibodies detect the same 68–64-kDa glycoprotein anchored to the plasma membrane by glycosyl-phosphatidylinositol linkage. We purified the MAX.1/11 antigen by immunofluorescence chromatography using monoclonal antibody MAX.11. The NH<sub>2</sub>-terminal amino acid sequence was determined and turned out to be identical to the NH<sub>2</sub>-terminal sequence of the membrane-bound carboxypeptidase M. By precipitation with antibodies MAX.1 and MAX.11, membrane preparations of macrophages and placental microvilli were almost completely depleted of enzyme activity, indicating that the two antibodies indeed recognize carboxypeptidase M. Immunoreactivity of both antibodies correlates with the reported tissue distribution of enzyme activity. Expression of carboxypeptidase M on mRNA level and enzymatic activity markedly increase during *in vitro* differentiation of monocytes, according to the described increase in MAX.1 and MAX.11 antigen expression. Moreover, *in vitro* differentiated macrophages show the highest specific activity yet described in any tissue. In addition, carboxypeptidase M expression could be detected in HL-60, U937, and THP-1 myeloid cell lines. Vitamin D<sub>3</sub>-induced monocyte differentiation resulted in an increased carboxypeptidase M expression in all three cell lines. Further studies are needed to elucidate the functional role of carboxypeptidase M during monocyte differentiation and activation.

Macrophages (MAC)<sup>1</sup> are believed to be the mature effector cells of the mononuclear phagocyte lineage. They participate in both specific and nonspecific immune responses (1) and play a

critical role in extracellular matrix remodeling and wound healing (2). Their differentiation from bone marrow precursors via circulating blood monocytes (MO) is a multistep process that is only partially understood (3, 4).

Peripheral blood MO provide the common source of MAC. Arbitrarily, or in response to so far unknown signals, they infiltrate tissues and body cavities. There, probably depending on the microenvironment and tissue-specific factors, they develop into different types of MAC with functional heterogeneity (4, 5). Under *in vitro* conditions a similar process of maturation from human blood MO to MAC, accompanied by specific changes of morphological, cytochemical, phenotypic, and functional properties can be observed (6, 7). Several cell surface antigens are commonly used as phenotypic markers; for example, the low affinity Fc receptor (CD16), the  $\alpha$ -chain of the vitronectin receptor (CD51), the transferrin receptor (CD71), and endoglin (CD105) are preferentially detectable on mature MAC, whereas the B148.4 antigen is mainly expressed on peripheral blood MO (8).

The two monoclonal antibodies MAX.1 and MAX.11 have originally been developed to detect lineage-restricted antigens specific for late differentiation stages of the mononuclear phagocyte lineage (9). During *in vitro* differentiation of human blood MO, surface expression of MAX.1 and MAX.11 antigen is increasing from low or undetectable levels on MO to a high expression level on MAC (after cultivation for 7 days in the presence of serum) (9). *In vivo*, MAC in serous cavities (e.g. pleural and peritoneal cavity and alveolar space) express MAX.1 and MAX.11 antigen in a heterogeneous pattern, whereas tissue MAC from various organs express them at a low level (10).

We purified MAX.1 and MAX.11 antigen by immunofluorescence chromatography from the supernatants of phosphatidylinositol-specific phospholipase C (PI-PLC)-treated, *in vitro* differentiated MAC. After NH<sub>2</sub>-terminal sequencing we identified carboxypeptidase M (CPM) as the antigen detected by both antibodies MAX.1 and MAX.11. CPM cleaves COOH-terminal arginine or lysine from peptides and proteins (11), and it is known to be anchored to the plasma membrane by glycosylphosphatidylinositol linkage (12). Because of its presence on plasma membranes and an optimal activity at neutral pH, CPM may participate in modulation of peptide hormone activity (e.g. kinins and enkephalin hexapeptides), or inactivation of peptides (e.g. kinins and anaphylatoxins) at local tissue sites (11, 13). Here we report a detailed study of CPM expression during *in vitro* differentiation of human blood MO. Furthermore, human myeloid leukemia cell lines appear to correspond with different immature stages of myeloid differentiation. The human myeloid leukemia cell lines THP-1, U937, and HL-60 differentiate into MO/MAC-like cells when they are treated with 1,25(OH)<sub>2</sub>D<sub>3</sub> (14–16), and this effect is mediated through the vitamin D<sub>3</sub> receptor (17). We investigated expres-

\* This work was supported by the Deutsche Forschungsgemeinschaft. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence and reprint requests should be addressed: Abt. f. Hämatologie und Internistische Onkologie, Klinikum der Universität Regensburg, 93042 Regensburg, Germany. Tel.: 49-941-944-7110; Fax: 49-941-944-7111; E-mail: Reinhard.Andreessen@klinik.uni-regensburg.de.

The abbreviations used are: MAC, macrophages; MO, monocytes; CD, cluster of differentiation; PI-PLC, phosphatidylinositol-specific phospholipase C; CPM, carboxypeptidase M; 1,25(OH)<sub>2</sub>D<sub>3</sub>, 1,25-dihydroxyvitamin D<sub>3</sub>; mAb, monoclonal antibody; PAGE, polyacrylamide gel electrophoresis; FACS, fluorescence-activated cell sorting; PBS, phosphate-buffered saline; NHS-LC-biotin, 6-(+)-biotinylamidohexanoic acid N-hydroxysuccinimide ester sodium salt; TSS, Tris/saline solution; NGF, N-glycosylase F; ECL, enhanced chemiluminescence; PCR, polymerase chain reaction; dansyl, 5-dimethylaminonaphthalene-1-sulfonyl.



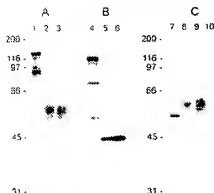


FIG. 1. Biochemical characterization of MAX.1 and MAX.11 antigen. SDS-PAGE analysis and ECL detection of immunoprecipitates obtained from lysates (A and B) or PI-PLC-supernatants of biotin surface-labeled *in vitro* differentiated MAC (C) with the following antibodies: CD11c mAb anti-Leu M5 (lanes 1, 4, and 7); mAb MAX.1 (lanes 2, 5, and 8); mAb MAX.11 (lanes 3, 6, and 9); and CD14 mAb RMO 82 (lane 10). B, immunoprecipitates were treated with *N*-glycosidase F before being subjected to SDS-PAGE.

Kit, Invitrogen, San Diego, CA). Partial DNA sequencing was done by dye decoupling terminator cycle sequencing (Applied Biosystems, Weiterstadt, Germany) according to the manufacturer's instructions, and sequences were analyzed on the Applied Biosystems DNA sequencing system (model 373A).

**Northern Analysis.**—Total RNA was isolated from different cell types by the guanidine thiocyanate/acid phenol method (28). The RNA (10  $\mu$ g/lane) was separated by electrophoresis on 1% agarose/formaldehyde gels (29) and transferred to nylon membranes (Magna NT, MSI, Westborough, MA). Hybridization (adapted from Refs. 29 and 30) was performed using a <sup>32</sup>P-labeled cDNA probe of the EcoRI restriction fragments of the cloned CPM PCR fragment (random primed DNA labeling kit, Boehringer Mannheim). To provide an internal control, membranes were reprobed with an oligonucleotide against 18S rRNA labeled by T4 kinase (5' end-labelling kit, Amersham). Autoradiography was performed at -70 °C, and blots were scanned with a Molecular Dynamics personal densitometer.

## RESULTS

**Biochemical Characterization of MAX.1 and MAX.11 Antigen: Immunoprecipitation, Deglycosylation, and PI-PLC Treatment.**—For a biochemical characterization of antigens detected by monoclonal antibodies MAX.1 and MAX.11, we surface-labeled *in vitro* differentiated MAC with NHSS-LC-biotin and immunoprecipitated antigens from the cell lysates. Both mAbs precipitated a protein with an apparent molecular mass of 58–64 kDa (Fig. 1A). Subsequent digestion of the immunoprecipitated material with *N*-glycosidase F resulted in a shift of apparent size to 46 kDa (Fig. 1B), indicating a carbohydrate content of about 25% by weight.

Since it was already shown that the antigen detected by MAX.1 can be cleaved from the cell membrane by PI-specific phospholipase C (31), immunoprecipitations were performed using supernatants of surface-labeled MAC after treatment with PI-PLC. The diffuse bands of the antigens precipitated by mAb MAX.1 and MAX.11 after PI-PLC treatment were identical to the bands obtained from cell lysates (Fig. 1C).

**Immunological Cross-reactivity of MAX.1 and MAX.11.**—Since the characteristics of antigens detected by mAb MAX.1 and MAX.11 were very similar during the experiments described above, we investigated the possibility that both antibodies detect the same antigen. Cell lysates from surface-labeled *in vitro* differentiated MAC were subjected to chromatography over small immunofluorescence columns (0.5 ml) of MAX.1 or MAX.11 covalently linked to Sepharose. After extensive washing to remove unspecifically bound material, eluates from both columns were reprecipitated with each of the two mAbs

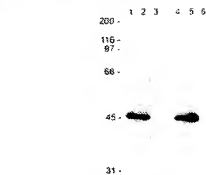


FIG. 2. Immunological cross-reactivity of mAbs MAX.1 and MAX.11. MAX.1 (lanes 1–3) and MAX.11 (lanes 4–5) immunofluorescence purification of lysates of biotin surface-labeled *in vitro* differentiated MAC and subsequent immunoprecipitation from the obtained eluates with the following antibodies: mAb MAX.1 (lanes 1 and 4); mAb MAX.11 (lanes 2 and 5); and CD11c mAb anti-Leu M5 (lanes 3 and 6). All immunoprecipitates were deglycosylated with *N*-glycosidase F before SDS-PAGE and ECL detection.

and an isotype control antibody. After enzymatic deglycosylation of the precipitated material the characteristic band of 46 kDa was detectable after precipitation with MAX.1 and MAX.11 but not with the isotype control antibody (Fig. 2), confirming that the antigens detected by MAX.1 and MAX.11 are identical.

**Purification and Identification of MAX.11 Antigen.**—*In vitro* differentiated MAC ( $8 \times 10^6$  cells) were exposed to PI-PLC to cause the release of glycosylphosphatidylinositol-anchored proteins into the supernatant. A smaller portion of cells ( $1 \times 10^6$  cells) was surface-labeled with biotin before PI-PLC treatment. Supernatants of labeled and unlabeled cells were combined and subjected to immunofluorescence column chromatography over MAX.11 covalently linked to Sepharose. After washing the material bound to the column was eluted in small fractions, and a small part of each fraction was analyzed on SDS-PAGE with subsequent immunoblotting and visualization of biotinylated protein by ECL. The fraction containing the immunopurified antigen was used for NH<sub>2</sub>-terminal amino acid sequencing, and the following sequence was obtained: Iaa-Asp-Phe-Asn-Tyr-X-X-Gln-Glu-Gly-Met-Glu-Ala. A search of the EMBL nucleic acid data bases revealed identity with the NH<sub>2</sub>-terminal sequenced amino acids of CPM (13).

**Immunoprecipitation and Enzyme Activity.**—To further confirm that CPM is identical with MAX.1/11 antigen, we measured CPM activity with dansyl-Ala-Arg as substrate in membrane preparations from *in vitro* differentiated MAC after immunoprecipitation with mAb MAX.1 or MAX.11 or an isotype control antibody. As a positive control with well known CPM activity we used membrane preparations of placental microvilli in a similar experiment. As shown in Table I, in comparison to the isotype control antibody (100%) the majority of enzymatic activity was precipitated by antibodies MAX.1 and MAX.11.

**Expression of CPM during Monocytic Differentiation.**—After it was proven that mAbs MAX.1 and MAX.11 detect CPM, we investigated the correlation of CPM surface expression with membrane-bound enzymatic activity and mRNA-expression during serum-induced *in vitro* differentiation of MO. Cultured MO were harvested at the indicated time points, and preparations of cells for flow cytometry, RNA extraction, and enzyme assay were done in parallel. As shown in Fig. 3, FACS analysis confirmed that the surface expression of CPM markedly increases during monocytic *in vitro* differentiation. A slightly different staining pattern of mAbs MAX.1 and MAX.11 was

TABLE I  
Immunoprecipitation of carboxypeptidase activity from membrane fractions of *in vitro* differentiated macrophages (MAK) or placental microvilli (PMV)

Antibodies	Source			
	MAK	PMV	MAK	PMV
Isotype control	805 ± 123*	100	210 ± 27	160
MAK.1	18 ± 40	6 ± 5	37 ± 34	12 ± 11
MAX.11	73 ± 8	9 ± 1	28 ± 25	9 ± 8

\* MSA, mean specific activity.

<sup>b</sup> Residual carboxypeptidase activity after immunoprecipitation with isotype control antibody (CD11c mAb), mAb MAX.1, and mAb MAX.11 was measured with dansyl-Ala-Arg as substrate. Data are given as mean specific activity (nmol/h/mg protein) ± S.E. from three independent experiments.

\* Mean  $\pm$  S.E. of control  $\pm$  S.E.



Fig. 3. Carboxypeptidase M surface expression. FACS analysis of cell surface expression for CPM during serum-induced differentiation of human M0 using mAbs MAX.1 (○) and MAX.11 (●). Freshly isolated M0 (day 0), M0 after adherence overnight (day 1), and at 1 (day 2), 3 (day 3), 4 (day 4), 7 (day 7) and 8 (day 8) days of culture in the presence of serum were used for FACS analysis. Monocytes were gated according to their forward and side scattering; >95% of gated cells were positive for CD14 and CPM expression in all cases. Data are expressed as mean channel fluorescence (MCF) of positive cells after subtraction of background obtained with an isotype control antibody from one representative experiment.

reproducible in all experiments. For both antibodies, the mean fluorescence, a parameter corresponding to the number of molecules/cell, strongly increased during 7 days of culture in the presence of serum.

In parallel, carboxypeptidase activity of cell membrane preparations was measured with dansyl-Ala-Arg as substrate at the same time points, using a placental membrane preparation as a positive control. The mean specific activity (nmol/h/mg) of CPM increased 30–40-fold during monocytic differentiation (Fig. 4), showing a similar behavior as the surface expression determined by FACS analysis.

To study the regulation of CPM expression at the RNA level, total cellular RNA was prepared at the same time points and hybridized with a CPM-probe. Total RNA prepared from placenta was used as a positive control, and lymphocyte RNA and fibroblast RNA were used as negative controls. As shown in Fig. 5, a strong induction of CPM mRNA expression was detectable during 7 days of culture in the presence of serum.

**Expression of CPM on Myeloid Cell Lines HL-60, U937, and THP-1**—In addition we studied expression of CPM on myeloid cell lines HL-60, U937, and THP-1, which appear to represent different immature stages of myeloid differentiation.

In promyeloid cell line HL-60 CPM surface expression was almost undetectable. Upon treatment with  $1,25(\text{OH})_2\text{D}_3$  a weak induction of CPM expression could be demonstrated in FACS analysis (Table II). The histiocytic cell line U937 showed a weak surface expression of CPM that could be slightly enhanced by treatment with  $1,25(\text{OH})_2\text{D}_3$ . In comparison with the other two cell lines the monoblastic cell line THP-1 showed a

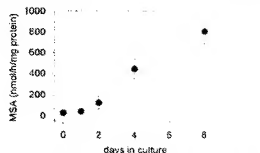


Fig. 4. Carboxypeptidase M enzymatic activity during monocytic differentiation *in vitro*. Membrane fractions were prepared from freshly isolated M0, M0 after adherence overnight, and M0 at 1, 3, and 7 days of culture in the presence of serum. Carboxypeptidase activity was measured with dansyl-Ala-Arg substrate; data are mean specific activity (MSA)  $\pm$  S.E. values ( $n = 4$ ).



Fig. 5. Northern blot analysis of carboxypeptidase M mRNA. Total RNAs were isolated from freshly isolated M0 (lanes 1, 2), M0 after adherence overnight (lanes 3, 4), and M0 at 1 (lanes 5, 6), 3 (lanes 7, 8), and 7 (lanes 9, 10) days of culture in the presence of serum, human lymphocytes (lanes 11), human placental microvilli (lanes 12), and human skin fibroblasts (lanes 13). The blot was sequentially hybridized with a <sup>32</sup>P-labeled CPM cDNA (panel A) and an 18S rRNA-specific oligonucleotide probe (panel B) to control for equivalent RNA loading.

TABLE II  
Carboxypeptidase M surface expression on myeloid cell lines

Myeloid Cell Lines	Treatment with $1,25(\text{OH})_2\text{D}_3$	Antibodies		
		CD14	MAX.1	MAX.11
		MCF	MCF	MCF
THP-1 <sup>a</sup>	+	4 ± 6	280 ± 100	310 ± 80
	-	400 ± 130	460 ± 50	510 ± 30
U937	+	0	24 ± 1	28 ± 4
	-	220 ± 190	43 ± 3	47 ± 6
HL-60	+	13 ± 1	14 ± 1	14 ± 2
	-	1600 ± 300	15 ± 1	53 ± 4

\* MCF, mean channel fluorescence.

<sup>a</sup> Cells ( $2 \times 10^6/\text{ml}$ ) were cultured for 72 h with or without  $1,25(\text{OH})_2\text{D}_3$  ( $10^{-7}$  M); antigen expression was determined by FACS using CD14 mAb Myc, mAb MAX.1, and mAb MAX.11.

<sup>b</sup> Data are given as mean channel fluorescence  $\pm$  S.E. values after subtraction of background obtained with isotype control antibody from at least three independent experiments.

high surface expression of CPM that markedly increased during the 3 days of  $1,25(\text{OH})_2\text{D}_3$ -treatment.

#### DISCUSSION

The expression and tissue distribution of the glycoproteins detected by mAbs MAX.1 and MAX.11 suggested that they might be of interest for the investigation of MAC differentiation and site-specific activation (9, 10). In this report, we describe the purification, characterization, and identification of the antigen detected by both mAbs. We used PI-PLC cleavage from the plasma membrane to purify the MAX.11 antigen by immunoaffinity chromatography. The quantity and purity of the eluted protein were sufficient for determination of  $\text{NH}_2$ -terminal amino acid sequence. A comparison of the determined

amino acid sequence with the published amino acid sequence of CPM (13) disclosed that the two molecules are identical. Results from biochemical characterization (58–64-kDa glycoprotein, anchoring to the plasma membrane by glycosylphosphatidylinositol linkage) matched the published data (11–13), and both antibodies were able to almost completely precipitate enzymatic activity from membrane preparations of *in vitro* differentiated MAC and placental microvilli as a control, too. Reactivity of MAX.1 and MAX.11 slightly differs in immunoperoxidase stainings of MAC from different body fluids (9, 10) and immunofluorescent flow cytometry during MO to MAC differentiation. Staining of CPM by mAb MAX.1 seems to be delayed in comparison with mAb MAX.11, suggesting that both antibodies might recognize differentially expressed epitopes on the CPM-molecule. Digestion of precipitated CPM with neuraminidase and  $\beta$ -galactosidase both lead to a marked shift of apparent size to 60 and 56 kDa, respectively (data not shown), indicating that the glycans linked to CPM are, at least in part, of the complex type. It is known that the expression of some glycosyltransferases responsible for glycan structure and composition can be regulated in a maturation-associated manner (33, 34). Therefore, a different glycosylation of CPM during the process of MO to MAC differentiation might be an explanation for the slightly different binding of MAX.1 and MAX.11.

Further evidence for the identity of CPM and MAX.1/11 antigen comes from the described similar tissue distributions of enzymatic activity and antigen expression. In addition to expression on MAC, MAX.1 and MAX.11 antigen could be detected in human kidney (glomerular mesangial cells), in human placenta (cytotrophoblastic and syncytial cells), in the upper gastrointestinal tract (subepithelial structures), and weakly in lymphoid organs (dendritic reticulum cells) (10). This staining pattern correlates with the published distribution of enzymatic activity of CPM. Membrane-bound carboxypeptidase activity is measurable in membrane fractions of various human tissues (placental microvilli, kidney, lung, and brain) (11, 12, 33–35) and in cultured human cells (forskolin fibroblasts, pulmonary arterial endothelial cells, lung fibroblasts, and alveolar type I cells) (33, 34).

Basic carboxypeptidases catalyze the removal of COOH-terminal basic amino acids arginine or lysine from peptides or proteins. The natural substrates appear to be peptide hormones like histone, enkephalin hexapeptides, and anaphylatoxins or proteins like creatine kinase (36). The removal of COOH-terminal arginine or lysine results in modulation or inactivation of peptide hormone activity and can also change the physical properties of proteins (36).

CPM is a membrane-bound ectopeptidase, present in many organs and tissues. It can participate in control of peptide hormone activity at the cell surface, degradation of extracellular proteins and peptides, and prohormone processing (36), but its regulation and actual functional role in different tissues are still unclear.

The release of L-arginine from peptides or proteins by CPM may increase local concentrations of this amino acid at the MAC cell surface. At sites of a high density of substrates, e.g., at inflammatory sites, CPM activity might influence L-arginine metabolism.

In the lung, CPM is present on alveolar type I pneumocytes (33, 34) and on alveolar MAC in yet lower concentrations (9, 10, 34). There is some evidence that CPM is involved in bradykinin metabolism. It may protect the alveolar surface from bradykinin-induced deleterious vascular responses (37–39).

In patients with extrinsic allergic bronchitis and active pulmonary sarcoidosis, CPM expression on alveolar MAC as detected by mAb MAX.1 is increased in comparison with healthy persons (10). Elevated serum levels of angiotensin-converting enzyme, a predominantly membrane-bound peptidyl dipeptidase that cleaves similar substrates as CPM, is used to diagnose and to assess response to treatment of sarcoidosis (40). CPM expression on alveolar MAC or its activity in the bronchoalveolar fluid could serve as additional markers for this disease. Further investigations should clarify the value of this parameter in diagnosing and monitoring active pulmonary sarcoidosis and other inflammatory lung diseases.

MAC from granulomas of patients with sarcoidosis synthesize the biological active hormonal form of vitamin D<sub>3</sub>, 1,25(OH)<sub>2</sub>D<sub>3</sub> (41). 1,25(OH)<sub>2</sub>D<sub>3</sub> may be involved in the stimulation of alveolar MAC to produce CPM. The possibility for a role of vitamin D<sub>3</sub> in regulation of CPM expression is supported by several additional observations: it is shown that, in the absence of serum proteins, MAX.1/CPM is expressed in a maturation-associated manner during 1,25(OH)<sub>2</sub>D<sub>3</sub>-induced *in vitro* maturation of MO to MAC (18). As demonstrated in this study, 1,25(OH)<sub>2</sub>D<sub>3</sub>-induced differentiation of human promonocytic leukemia cell lines THP-1 and U937 is accompanied by an increased expression of CPM. Also in HL-60 cells, which are thought to represent the earliest differentiation stage of the three cell lines tested, CPM expression slightly increased during 1,25(OH)<sub>2</sub>D<sub>3</sub> treatment.

In renal allograft rejection T-cell activation is required to cause CPM expression in local MAC, since the CPM expression is undetectable in patients treated with Cyclosporin A (10). These results indicate that T-lymphokines may have a possible regulatory role in the expression induction of CPM. Recent observations of our group suggest that tumor cells modulate MAC differentiation and are capable of suppressing CPM expression on MAC associated with solid tumors.<sup>3</sup>

To clarify whether T-cell-derived or tumor-derived factors play an active role in modulation of CPM expression further studies on gene regulation are required.

The functional role of CPM in MAC will depend on its tissue localization and the occurrence of its natural substrates. The present available data on possible substrates are still incomplete. Peptide hormone processing includes complex regulatory mechanisms. Its course and its role in tissue injury and inflammation as well as the effects of peptide hormones and metabolites on cells of the immune system have not yet been characterized.

As described previously (9, 10), CPM expression, detected by mAb MAX.1 and MAX.11, is restricted to the MO/MAC lineage within hemopoietic cells. Besides other peptidases (neutral endopeptidase (CD10), aminopeptidase N (CD13), and dipeptidyl-peptidase IV (CD26)) carboxypeptidase M should be assigned to a cluster differentiation (CD) group.

**Acknowledgments**—We thank Dr. Frank Enmerich for providing sufficient amounts of MAX.1 and MAX.11 hybridoma supernatants, Dr. Rainer Deutzman for performing the N<sub>175</sub>-terminal sequencing, Anne Petry-Krieger for cDNA sequencing, and Dr. Werner Fals for helpful advice and discussions. The excellent technical assistance of Lucia Schwarzfischer-Pfeilthöfer and Ute Ackermann is greatly appreciated.

#### REFERENCES

1. Van Furth, R. (1992) *Immunobiology* 161, 175–185.
2. Rappole, D. A., Mark, A., Banda, M. J., and Werb, Z. (1988) *Science* 241, 709–712.
3. Metcalf, D. (1991) *Science* 254, 520–523.
4. Rutherford, M. S., Witzel, A., and Schenk, L. B. (1993) *J. Leukocyt. Biol.* 55.

<sup>3</sup> M. Rehl, S. W. Krause, M. Kreutz, and R. Andreessen, unpublished observations.

<sup>4</sup> A. Konur, M. Kreutz, R. Knüchel, and R. Andreessen, submitted for publication.

- 602-618
5. Gordon, S., Fraser, I., Nath, D., Hughes, D., and Clarke, S. (1992) *Curr. Opin. Immunol.* **4**, 25-32.
6. Messan, R. A. (1985) *Am. J. Physiol.* **111**, 331-340.
7. Andresson, R., Picht, J., and Lohr, G. W. (1983) *J. Immunol. Methods* **56**, 295-304.
8. Andresson, R., and Kruetz, M. (1994) *Ernst Schering Research Foundation Workshop 11: Cell Culture in Pharmaceutical Research* (Fasentin, E., and Graf, H., eds) pp. 9-37. Springer-Verlag Berlin, Heidelberg.
9. Andresson, R., Bruns, K. J., Osterholz, J., and Emmerich, F. (1998) *Blood* **67**, 1237-1244.
10. Andresson, R., Gadd, S., Costabel, U., Leser, H. G., Speth, V., Cuenik, B., and Atkins, R. C. (1998) *Gell Tissue Res.* **285**, 271-279.
11. Skidgel, R. A., Davis, R. M., and Tan, P. (1989) *J. Biol. Chem.* **264**, 2235-2241.
12. Deddish, P. A., Skidgel, R. A., Kriha, V. B., Li, X.-Y., Becker, R. P., and Erdos, E. G. (1990) *J. Biol. Chem.* **265**, 15083-15088.
13. Tan, P., Chen, S. J., Steiner, D. F., Schilling, J. W., and Skidgel, R. A. (1989) *J. Biol. Chem.* **264**, 13185-13179.
14. Tanaka, H., Abe, E., Miyaura, C., Karibayashi, T., Konno, K., Nishii, Y., and Suda, T. (1992) *Biochem. J.* **204**, 713-719.
15. Bar-Shavit, Z., Telatbaum, S. L., Raitera, P., Hall, A., Pegg, L. E., Trist, J., and Kahn, A. J. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5907-5911.
16. Olsson, I., Dahlberg, U., Ivrek, I., and Nilsson, K. (1983) *Cancer Res.* **43**, 5852-5857.
17. Mangelstorf, D. J., Konfler, H. P., Doudman, C. A., Pike, J. W., and Hausader, M. R. (1984) *J. Cell Biol.* **98**, 391-398.
18. Kruetz, M., and Andresson, R. (1990) *Blood* **76**, 2457-2461.
19. Andresson, R., Bruggen, W., Scheibbogen, C., Kruetz, M., Leser, H. G., Rehm, A., and Lohr, G. W. (1990) *J. Leukocyte Biol.* **47**, 489-497.
20. Cole, S. R., Ashman, L. K., and Ry, P. L. (1987) *Mol. Immunol.* **24**, 699-705.
21. Krieger, T., Arnst, S., Malarkannan, S., Pannestiel, M., and Hoersch, D. (1992) *Anal. Biochem.* **204**, 220-226.
22. Shapiro, A. L., Vinuela, E., and Maizel, J. V. (1967) *Biochem. Biophys. Res. Commun.* **129**, 155-164.
23. Lammli, U. K. (1970) *Nature* **227**, 680-685.
24. Wessel, D., and Flügge, U. I. (1984) *Anal. Biochem.* **138**, 141-143.
25. Tewks, H., Stachetia, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350-4354.
26. Skidgel, R. A. (1991) *Methods Neurosci.* **6**, 373-385.
27. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallik, A. K., Gartner, F. H., Fawcett, M. D., Fajimoto, E. K., Goeke, N. M., Olson, B. J., and Jenk, D. C. (1985) *Anal. Biochem.* **156**, 76-85.
28. Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162**, 156-159.
29. Krause, S. W., Kruetz, M., Zenke, G., and Andresson, R. (1992) *Ann. Hematol.* **64**, 190-195.
30. Church, G. M., and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1591-1595.
31. Van der Schoot, C. E., Huizinga, T. W. J., Gadd, S., Majid, O., Wijman, R., Knapp, W., and von dem Borne, A. E. G. K. (1989) in *Leukocyte Typing IV* (Knapp, W., Dörken, B., Gilks, W. E., Rieber, E. P., Schmidt, R. E., Siena, H., and von dem Borne, A. E. G. K., eds), pp. 897-891. Oxford Press, Oxford.
32. Mizoguchi, A., Takasaki, S., Maeda, S., and Kobata, A. (1985) *J. Biol. Chem.* **260**, 1949-1957.
33. Skidgel, R. A., Johnson, A. R., and Erdos, E. G. (1994) *Biochem. Pharmacol.* **38**, 3471-3478.
34. Nagase, A., Abe, M., Becker, R. P., Deddish, P. A., Skidgel, R. A., and Erdos, E. G. (1993) *Am. J. Respir. Cell Mol. Biol.* **9**, 221-229.
35. Nagase, A., Deddish, P. A., Becker, R. P., Anderson, C. H., Abe, M., Tan, P., Skidgel, R. A., and Erdos, E. G. (1992) *J. Neurochem.* **59**, 2201-2212.
36. Skidgel, R. A. (1988) *Trends Pharmacol. Sci.* **9**, 299-304.
37. Chidambala, V., Skidgel, R. A., Kruetz, E. J., and Murias, C. G. (1991) *Am. Rev. Respir. Dis.* **144**, 859-874.
38. Desmanes, N., Lookhart, A., Lacroix, H., and Dunbar, D. J. (1992) *Am. J. Respir. Cell Mol. Biol.* **7**, 477-484.
39. Dragovic, T., Igle, R., Erdos, E. G., and Rabbit, S. F. (1993) *Am. Rev. Respir. Dis.* **147**, 1491-1498.
40. Lieberman, J. (1989) *Clin. Lab. Med.* **9**, 745-755.
41. Adams, J. S., Singer, F. R., O'Connell, M. A., Sharara, O. P., Hayes, M. J., Younes, F., and Haisch, M. F. (1985) *J. Clin. Endocrinol. & Metab.* **60**, 960-966.

# Opposing Effects of Ras on p53: Transcriptional Activation of *mdm2* and Induction of p19<sup>ARF</sup>

Stefan Ries,<sup>1,2</sup> Carola Biederer,<sup>1,2</sup>  
Douglas Woods,<sup>1</sup> Ohad Shifman,<sup>1</sup>  
Senji Shirasawa,<sup>3</sup> Takehiko Sasazuki,<sup>4</sup>  
Martin McMahon,<sup>5</sup> Moshe Oren,<sup>1</sup>  
and Frank McCormick<sup>1,2</sup>

<sup>1</sup>University of California, San Francisco  
Cancer Research Institute  
San Francisco, California 94115

<sup>2</sup>Department of Molecular Cell Biology  
Weizmann Institute of Science  
Rehovot 76100  
Israel  
<sup>3</sup>Kyushu University  
Fukuoka, Japan

## Summary

Mdm2 acts as a major regulator of the tumor suppressor p53 by targeting its destruction. Here, we show that the *mdm2* gene is also regulated by the Ras-driven Raf/MEK/MAP kinase pathway, in a p53-independent manner. Mdm2 induced by activated Raf degrades p53 in the absence of the Mdm2 inhibitor p19<sup>ARF</sup>. This regulatory pathway accounts for the observation that cells transformed by oncogenic Ras are more resistant to p53-dependent apoptosis following exposure to DNA damage. Activation of the Ras-induced Raf/MEK/MAP kinase may therefore play a key role in suppressing p53 during tumor development and treatment. In primary cells, Raf also activates the Mdm2 inhibitor p19<sup>ARF</sup>. Levels of p53 are therefore determined by opposing effects of Raf-induced p19<sup>ARF</sup> and Mdm2.

## Introduction

Mdm2 was originally identified as an amplified gene in a spontaneously transformed derivative of BALB/c cell line 3T3 DM, which caused tumors when injected into nude mice (Fakhrazadeh et al., 1991). A possible mechanism for the transforming properties of *mdm2* has been provided by reports demonstrating that Mdm2 is a major regulator of the tumor suppressor p53. It binds directly to p53 and inhibits its transcriptional activity (Mornand et al., 1992; Oliner et al., 1992, 1993). Mdm2 is a transcriptional target of p53 (Barak et al., 1993; Wu et al., 1993; Leng et al., 1995). p53-responsive elements have been identified in the intronic promoter of the *mdm2* gene, and interaction of p53 with these sites has been well documented (Juven et al., 1993; Barak et al., 1994; Zauberman et al., 1995). Induction of *mdm2* transcription by p53 establishes a negative feedback loop, in which p53 itself initiates its own destruction (Picksley and Lane, 1993).

Mdm2 may have additional functions that are not directly related to p53. It affects cell growth in a p53-independent manner, possibly through interactions with

pRb (Xiao et al., 1995) or through interaction with the E2F/DP1 complex (Martin et al., 1995). These functions of Mdm2 are less well characterized than its role in p53 regulation. Indeed, the lethal effects of disrupting *mdm2* genes in vivo are rescued by disrupting p53, suggesting that p53 regulation is the major function of this protein at least during early development (Jones et al., 1995; Montes de Oca Luna et al., 1995).

Mdm2 binds to p19<sup>ARF</sup> and is inhibited by this interaction (Kamijo et al., 1998; Pomerantz et al., 1998). An attractive model has been presented recently in which p19<sup>ARF</sup> binds to Mdm2, sequesters Mdm2 in nucleolar structures, and allows accumulation of p53 (Tao and Levine, 1999; Sherr and Weber, 2000). On the other hand, p19<sup>ARF</sup> directly inhibits Mdm2 ubiquitin ligase activity, suggesting a more direct role in Mdm2 regulation (Honda and Yasuda, 1999). p14<sup>ARF</sup>, the human homolog of p19<sup>ARF</sup>, is induced by E2F, myc and Ras, and thus provides a possible link from mitogenic signaling pathways to p53 induction (Bates et al., 1998; Palmero et al., 1998; Zindy et al., 1998). The Ras-regulated Raf/MEK/ERK kinase pathway has been reported to activate CDK4/cyclin D kinases, thereby phosphorylating pRb, which in turn leads to release of E2F-1 (Albanese et al., 1995; Peepers et al., 1997). The p14<sup>ARF</sup> promoter contains several E2F-1 binding sites and its activity was shown to be enhanced by overexpression of E2F-1 (Bates et al., 1998). The Ras/Raf/MEK/MAP kinase pathway can therefore lead indirectly to accumulation of p14<sup>ARF</sup> and inhibition of Mdm2 activity.

Mdm2 expression is often increased following mitogenic activation. Treatment of cells in culture with basic FGF increases levels of Mdm2 protein and cells constitutively exposed to a basic FGF autocrine loop are more refractory to killing by cisplatin, which to a large extent occurs through p53-mediated apoptosis (Shaulian et al., 1997). More recently, a screen for transcripts that accumulate in cells harboring a chimeric M-CSF/PDGF receptor identified *mdm2* as an immediate early gene (Fambrough et al., 1999). In addition, we have observed high levels of Mdm2 protein expression in human tumor cell lines that have little, if any, functional p53 (Ries et al., 2000). These data suggest regulation of *mdm2* expression by p53-independent pathways triggered by growth factors.

A common feature of signaling by diverse growth factors is the activation of the Ras/Raf/MEK/MAP kinase pathway. We therefore explored the possibility that this pathway is responsible for *mdm2* induction. We report that the *mdm2* promoter is indeed a target of the Ras/Raf/MEK/MAP kinase pathway. Hence, activation of Ras during normal cell signaling or through mutation in neoplastic transformation, can suppress p53 and thus facilitate cell proliferation and survival.

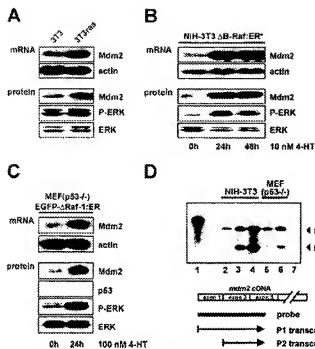
## Results

### Mdm2 Expression Is Regulated by the Ras/Raf Pathway

Mdm2 protein can be induced by exposure of cells to basic FGF (Shaulian et al., 1997) or IGF-1 (Leri et al., 1999). Likewise, *mdm2* mRNA accumulates upon activation of the PDGF receptor (Fambrough et al., 1999).

<sup>5</sup> To whom correspondence should be addressed (e-mail: mcmccormick@cc.ucsf.edu).

<sup>2</sup> These authors contributed equally to this work.



each RNA was subjected to RNase protection analysis. Lanes 1 and 7 contain the probe alone and the control reaction consisting of 10  $\mu$ g of yeast tRNA, respectively. Composition and sizes (in nucleotides) of potential RNase-resistant fragments are indicated in the schematic representation.

Since Ras is downstream of all these mitogens, we tested whether the Ras pathway is responsible for *mdm2* induction. NIH-3T3 cells expressing an activated Ras allele, H-Ras (G12V), were analyzed for *mdm2* mRNA and protein expression. Figure 1A shows that expression of constitutively active Ras resulted in increased levels of *mdm2* mRNA and protein.

Ras targets several distinct downstream effectors, of which the best characterized are Raf kinase, PI3-kinase, and RalGDS (Marshall, 1995; Katz and McCormick, 1997; Rodríguez-Viciana et al., 1997). We examined the effects of a conditionally active Raf kinase on cellular Mdm2 levels. This conditional Raf kinase consists of the kinase domain of B-Raf fused to a modified form of the hormone binding domain of the mouse estrogen receptor, which renders the receptor insensitive to estrogen, but still sensitive for the estrogen analog 4-hydroxy-tamoxifen (4-HT) (Samuels et al., 1993; Woods et al., 1997). NIH3T3 cells stably expressing the  $\Delta$ B-Raf:ER<sup>+</sup> were treated with 10 nM 4-HT (Woods et al., 1997). Addition of 4-HT resulted in remarkable MAPK activation and accumulation of high levels of *mdm2* mRNA and protein (Figure 1B). To rule out any contribution of p53 to the increased Mdm2 levels after Raf activation, we assessed Mdm2 expression in p53<sup>-/-</sup> mouse embryonic fibroblasts (MEFs) stably expressing EGFP- $\Delta$ Raf1:ER (Woods et al., 1997). Figure 1C shows increased transcription and accumulation of Mdm2 protein in response to Raf/MEK/MAPK activation occurred also in p53<sup>-/-</sup> fibroblasts. In a time course experiment using NIH3T3( $\Delta$ B-Raf:ER<sup>+</sup>) cells, elevated *mdm2* mRNA levels can be detected as early as 4 hr after Raf activation and increase steadily (data not shown). Thus, the Ras/Raf/MEK/MAP kinase pathway induces increased expression of *mdm2* mRNA and Mdm2 protein in a p53-independent manner.

#### Figure 1. Induction of *mdm2* Transcription by Activated Ras and Raf

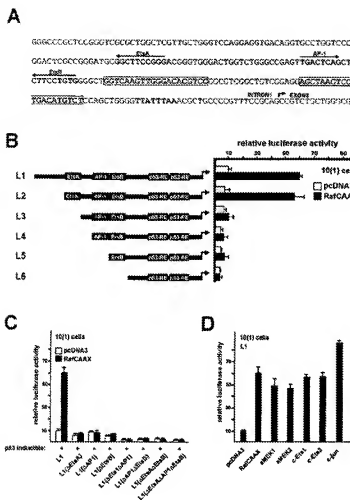
(A) *Mdm2* transcriptional induction by constitutively active H-Ras(G12V) in NIH-3T3 cells. NIH-3T3 cells were infected with retroviruses encoding H-Ras(G12V) or empty vector. After selection with G418, total cellular RNA was prepared and subjected to Northern blot analysis using a 1.5 kb fragment spanning the coding region of *mdm2* as a probe. For Western blot analysis, cells were directly lysed with Laemmli buffer, and equal amounts of total protein were subjected to immunoblotting. (B) Activation of Raf is sufficient to induce *mdm2* transcription in NIH-3T3 cells. NIH-3T3 cells expressing a 4-Hydroxy-Tamoxifen (4-HT) inducible  $\Delta$ B-Raf:ER<sup>+</sup> construct were treated with 10 nM 4-HT, total cellular RNA and protein harvested at the times indicated and subjected to Northern blot analysis or Western blot analysis. (C) Transcriptional induction of *mdm2* by Raf is independent of p53. p53<sup>-/-</sup> mouse embryonic fibroblasts (MEF) harboring the 4-HT inducible EGFP- $\Delta$ Raf1:ER construct were treated with 100 nM 4-HT, and total cellular RNA and protein was harvested after 24 hr. (D) Transcription by activated Raf initiates mainly from the internal *mdm2* promoter (P2). Total cellular RNA was extracted from the cell lines as indicated (top). Ten micrograms of

The murine *mdm2* gene has two promoters, an internal promoter (P2), which responds directly to p53 activation, and an upstream constitutive promoter (P1), which is not affected by p53 (Barak et al., 1994). The cDNA riboprobe used for RNase protection assay spans exon 1, exon 2, and part of exon 3 (Figure 1D), allowing discrimination between P1 and P2 transcripts. Transcription from the P2 promoter was dramatically induced in response to activation of the Raf kinase, whereas a more moderate induction occurred with transcription from P1 (Figure 1D, lanes 2, 3, and 4). Induction of P2 transcription occurs in a p53-independent manner as p53-deficient MEFs show similar enhanced transcription arising from P2 (Figure 1D, lanes 5 and 6). Note that only P1 transcripts are detectable in control NIH-3T3 cells.

#### Analysis of Ras-Responsive Elements within the *mdm2* P2 Promoter

To determine the role of Raf activation on basal *mdm2* transcription, the effect of a constitutively activated form of Raf (Raf-CAAX) on *mdm2* promoter activity was examined in p53-deficient 10T1 cells (Stokoe et al., 1994). Raf-CAAX expression induced *mdm2* P2 promoter activity 5- to 6-fold (Figure 2B). Sequence analysis of the *mdm2* P2 promoter revealed the existence of binding sites for transcription factors of the AP-1 and Ets family (Figure 2A), which have been shown to be responsive to ERK activation in other genes. To determine the region of the *mdm2* P2 promoter required for regulation by Raf, several *mdm2* P2 5' promoter deletions were generated and transfected in the presence of Raf-CAAX into 10T1 cells. Figure 2B shows that the upstream Ets site within the P2 promoter (designated EtsA in Figure 2B) of the *mdm2* gene is important for the Raf-induced





**Figure 2. Ets/AP-1 Motifs in the Intronic P2 Promoter Mediate Transcription of the *mdm2* Gene upon Raf Activation**

(A) DNA sequence of the intronic *mdm2* promoter (P2). The Ets and AP-1 binding sites are indicated by arrows, the p53-responsive elements are boxed, and the TATA box is marked in bold.

(B) Various 5' deletions of the *mdm2* promoter were cloned in front of the firefly luciferase gene. 10(1) cells were cotransfected with 1  $\mu$ g of each luciferase reporter construct and either 100 ng of a vector expressing CAX-tagged Raf or the parental control vector (pcDNA3). Luciferase activity was determined using a luminometer.

(C) Using in vitro mutagenesis, Ets and AP-1 sites of the *mdm2* promoter luciferase construct L1 as in (B) were selectively mutated. Luciferase activity of the mutated promoter constructs in 10(1) cells was determined. All promoter constructs were responsive to co-transfected p53 (indicated as +).

(D) 10<sup>5</sup> cells were transiently transfected with 1 µg of the *midm2* promoter luciferase construct L1 combined with 100 ng of plasmids directing expression of either constitutively active MEK1, c-Ets-1, c-Ets-2, or c-jun. Luciferase activity was determined.

transcription activation in a transient assay. To assess the importance of the more downstream Ets (designated EtsB) and AP-1 sites in the *mdm2* P2 promoter in response to activated Raf, single point mutations were introduced into the full-length *mdm2* P2 promoter construct 1 by site-directed mutagenesis. All promoter constructs that contain mutations in one of the Ets sites or the AP-1 site are nonresponsive to activated Raf anymore (Figure 2C). Furthermore, combinations of mutations in the Ets sites and the AP-1 site significantly increase basal *mdm2* P2 promoter activity (Figure 2C). To exclude the occurrence of potential nonspecific mutations during the generation of these constructs and to address whether the Raf kinase and p53 responsive elements are distinguishable, p53 was cotransfected with the mutated P2 promoter constructs in parallel experiments. All mutated *mdm2* promoter constructs showed a strong induction by p53 (~10-fold) similar to that of the parental 1 *mdm2* promoter construct, demonstrating a clearly distinct regulation of the P2 promoter by either Raf or p53 (data not shown). Overexpression of constitutively active Raf, p34<sup>RAF</sup> and c-Jun, but not of c-Fos, c-Ets, and c-Jun<sup>1</sup>, which are the downstream targets of Ras and Raf, led to an approximately 5- to 8-fold induction of *mdm2* P2 promoter activity, respectively (Figure 2D). To determine whether the

*mdm2* AP-1 and Ets sites were capable of binding the respective transcription factors, nuclear extracts were prepared from NIH-3T3 cells expressing *B-Raf:ER* at various times after stimulation with 10 nM 4-HT. The *mdm2* AP-1 site in probe GS1 bound a complex in NIH-3T3 nuclear extracts, the formation of which was increased strongly after activation of *Raf* (Figure 3A). This complex was competed by 50-fold molar excess cold cognate competitor oligonucleotide or unlabeled GS1 probe but not by mutant AP-1 sequences or unlabeled probe GS2 (Figure 3A). When the labeled upstream *mdm2* Ets site (EtsA) was used as probe in electrophoretic mobility shift assays, formation of a complex could be detected with intensity remaining constant after *Raf* activation. These studies demonstrate that transcription factors bind to the AP-1 and Ets elements in the *mdm2* promoter.

**Raf Regulates p53 Levels through Its Effects on Mdm2** Induction of Mdm2 protein expression is expected to decrease levels of p53, since Mdm2 effectively promotes p53 degradation (Haupt et al., 1997; Kubbutat et al., 1997). To test this prediction, Raf kinase was activated in human tumor cells expressing mutant p53, and levels of Mdm2 and p53 proteins were analyzed by Western blotting. DKO4 cells were used, in which mutant

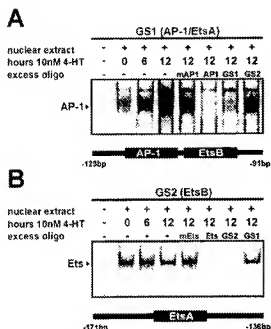


Figure 3. Transcription Factors Bind to the AP-1 and Ets Elements in the *mdm2* Promoter

(A) Nuclear extracts were prepared from NIH-3T3 cells expressing EGFP-ΔRaf-1:ER at the indicated time points after stimulation with 4-HT. EMSAs were performed with <sup>32</sup>P-labeled GS1-oligonucleotide comprising the AP-1/Ets element. In competition experiments, 50-fold molar excess of cold AP-1 or mutated AP-1 consensus oligonucleotide (AP1 and mAP1, respectively) was added to the binding reaction. (B) EMSAs were carried out with <sup>32</sup>P-labeled GS2-oligonucleotide comprising the Ets site of the *mdm2* promoter.

Ras had been deleted by homologous recombination (Shirasawa et al., 1993), and a conditionally active Raf allele (EGFP-ΔRaf-1:ER) was stably expressed in these cells (Woods et al., 1997). Raf kinase was turned on by addition of 4-HT, and Mdm2 levels accumulated (Figure 4A). p53 is transcriptionally inactive in these cells (Esteller et al., 2000). Hence, these data further confirm that Raf-dependent *mdm2* transcription is independent of p53 transcriptional activity. Importantly, increased Raf activity led to a dramatic decrease of p53 protein levels (Figure 4A). It is important to note that DKO4 cells, despite harboring a mutated p53 gene, do not express p14<sup>ARF</sup> due to hypermethylation of the p14<sup>ARF</sup> promoter (Esteller et al., 2000).

SW480 cells express an activated Ras allele and a mutant form of p53 that is transcriptionally inactive (Sharma et al., 1993). Treatment of these cells with the MEK inhibitor U0126 resulted in dose-dependent decrease in Mdm2 protein expression (Figure 4B) consistent with reduced MAPK activation, showing that sustained activity of the Ras/Raf/MEK/MAP kinase pathway is necessary for high levels of Mdm2.

To examine whether decrease of Mdm2 protein expression results from reduced transcriptional activity of the *mdm2* promoter upon MAPK inhibition, we measured *mdm2* P2 promoter activity. SW480 cells were transfected with either the *mdm2* P2 promoter construct L1, containing all the Raf-responsive elements, or the

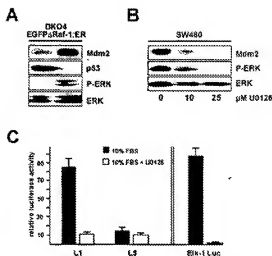


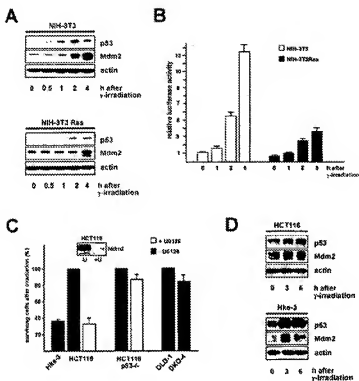
Figure 4. The Ras/Raf/MEK/MAPK Pathway Has an Impact on Mdm2 Expression Levels in Colon Cancer

(A) DKO4 cells expressing an EGFP-ΔRaf-1:ER construct were treated with 100 nM 4-HT. Twenty-four hours later, total cellular lysates were prepared. Equal amounts of proteins were separated by SDS-PAGE followed by immunoblotting with antibodies against Mdm2, p53, phosphorylated ERK1/2, or total ERK. (B) SW480 cells were cultured 48 hr in the presence or absence of the MEK inhibitor U0126 and 10% fetal bovine serum. Equal amounts of proteins were electrophoretically separated prior to immunoblotting with Mdm2, phosphorylated ERK1/2, or total ERK antisera. (C) SW480 cells were transfected with either 1 μg of the *mdm2* promoter luciferase construct L1 or the further 5'-deleted promoter construct L5, which lacks the Ras responsive elements. Transfected cells were incubated 48 hr in the presence or absence of 25 μM U0126 prior to measurement of luciferase activity. To monitor efficacy of the MEK inhibitor, U0126 Elk-1 activity was determined using a commercially available Elk-1 luciferase reporter system (right panel).

5' deleted L5 *mdm2* P2 promoter construct lacking the Ets and AP-1 binding motifs. Immediately after transfection, one fraction of the cells was incubated with 25 μM U0126 for 48 hr, while the other fraction of cells was incubated in medium without the MEK inhibitor. As displayed in Figure 4C, an approximately 7-fold decrease of *mdm2* P2 promoter activity was observed upon U0126 treatment only in cells transfected with the Raf responsive *mdm2* promoter construct L1. Luciferase activity of cells containing the *mdm2* promoter construct L5 was significantly lower and did not change upon addition of MEK inhibitor U0126.

#### Ras Mutations Lead to Attenuation of the p53 Response and Increased Survival Rate upon γ-Irradiation

To investigate whether the increased Mdm2 protein levels due to constitutive activation of Ras can prevent or attenuate p53 induction in response to cellular stress, we γ-irradiated NIH-3T3 cells or NIH-3T3 cells stably expressing a vector encoding H-Ras (G12V). Figure 5A shows the time course of p53 induction in both cell lines. In the parental NIH-3T3 cells, p53 protein was dramatically induced after γ-irradiation. Increased p53 levels are already detectable after 30 min and reach a



**Figure 5.** Kinetics of p53 induction upon  $\gamma$ -irradiation and radiation survival are affected by the Ras/Raf/MEK/MAPK pathway. (A) Logarithmically growing cultures of NIH-3T3 cells stably transfected with oncogenic H-Ras(G12V) or vector alone were subjected to  $\gamma$ -irradiation at a dosage of 3 Gy. Proteins were electrophoretically separated, immunoblotted, and immunoblotted with antibodies specific to Mdm2, p53, or  $\beta$ -actin. Lysates (NIH-3T3 and NIH-3T3 Ras) were run on the same gel. Exposure time is identical for each protein shown, allowing direct comparison of signal intensity. (B)  $1 \times 10^5$  NIH-3T3 cells and Ras-transformed NIH-3T3 cells, respectively, were seeded in 6-well plates. The next day, cells were transfected with 1  $\mu$ g p53-responsive luciferase reporter and exposed to ionizing radiation at a dosage of 3 Gy. Luciferase activity was determined using a luminometer. (C) Clonal survival assays were performed using the human colon cancer cell lines HCT116 and DLD-1, and their isogenic derivatives Hke3 and DLD-1 p53<sup>-/-</sup>, and DKO4, respectively. Cells were irradiated at a dosage of 5 Gy in a <sup>137</sup>Cs source. Survival rate for HCT116, HCT116 p53<sup>-/-</sup>, and DLD-1, respectively, was set 100% in each set of comparisons. (D) Time course of Mdm2 and p53 protein levels in HCT116 cells and its derivative Hke-3 after  $\gamma$ -irradiation (3 Gy). All lysates were run on the same gel. Exposure time is identical for each protein shown, allowing direct comparison of signal intensity.

maximum at 2 hr post irradiation. After 4 hr, p53 protein levels decreased but were still above the basal level in nonirradiated cells. Expression of Mdm2 protein in NIH-3T3 cells was strongly induced beginning 1 hr after irradiation of the cells and peaked at 4 hr. In contrast,  $\gamma$ -irradiation of NIH-3T3 cells harboring an activated Ras led to only a moderate increase of p53 protein. Furthermore, the onset of p53 induction in these cells is delayed compared to the parental NIH-3T3 cell line (1 hr in Ras-transformed NIH-3T3 cells versus 30 min in NIH-3T3 cells). While Ras-transformed NIH-3T3 cells have already higher basal Mdm2 protein levels (originating from the induced transcription of the *mdm2* P2 promoter by activated Ras), there was no significant increase of Mdm2 protein up to 4 hr after  $\gamma$ -irradiation. Thus, NIH-3T3 cells harboring an activated form of Ras are capable of attenuating the p53 response upon  $\gamma$ -irradiation, which is consistent with their increased basal Mdm2 protein expression.

Next, we wished to investigate the activity of the induced p53 in response to irradiation. We examined p53 activity after  $\gamma$ -irradiation in NIH-3T3 Ras cells using a luciferase assay. For this purpose, we transfected a synthetic p53 luciferase reporter construct into NIH-3T3 and Ras-transformed NIH-3T3 cells. Twenty-four hours after transfection, cells were irradiated at a dosage of 3 Gy and luciferase activity was determined at the indicated time points. Figure 5B displays the time course of p53 activation after  $\gamma$ -irradiation. While NIH-3T3 cells reveal a 12-fold induction of p53 transcriptional activity 4 hr after irradiation, their Ras-transformed counterparts show only a modest p53 response (4-fold). Thus, Ras overexpression attenuates not only p53 accumulation but also p53 activation in response to DNA damage.

This attenuation of p53 might explain the higher resistance of transformed NIH-3T3 cells toward  $\gamma$ -irradiation as determined in clonal survival assays (data not shown) (Sklar, 1988). Because Ras activation targets many downstream effectors (Katz and McCormick, 1997), we examined whether increased Mdm2 protein levels could mimic Ras-mediated radioresistance. Therefore, we established a stable NIH-3T3 cell line expressing Mdm2 protein under control of a mitopristone-inducible promoter (NIH-3T3-imdm2). Mitopristone-treated NIH-3T3-imdm2 cells and their parental counterpart, stably expressing the empty vector, were irradiated and subsequently, clonal survival assays were performed. Mdm2 expression partially rescues NIH-3T3 cells from irradiation-induced apoptosis (92% survivors in Mdm2-expressing NIH-3T3 versus 78% at 3 Gy; 49% versus 29% at 5 Gy).

To determine whether oncogenic Ras mutations confer radioresistance in human tumor cells, we performed clonal survival assays with the human colon cancer cell line HCT116 (wild-type p53, mutant Ras) and its derivative Hke-3, in which the mutant Ras has been deleted by homologous recombination (Shirasawa et al., 1993). Figure 5 shows Hke-3 express less Mdm2 protein and more p53 than HCT116 cells, consistent with a role for Ras in this pathway. p53 accumulated to significantly higher levels in Hke-3 cells after  $\gamma$ -irradiation, presumably because of lower levels of Mdm2 (Figure 5D), and radioresistance of these cells increased dramatically (80% survivors versus 28% survivors at 5 Gy; similar results were obtained at a dosage of 1 Gy). Furthermore, inhibition of MEK in the parental HCT116 cell line through addition of U0126 resulted in a similar radioresistant phenotype as in Hke-3 cells (Figure 5D). Treatment of

the p53-deficient isogenic cell line HCT116 p53<sup>-/-</sup> (Bunz et al., 1998) with the MEK inhibitor U0125 revealed no significant difference in radiosensitivity, showing that Ras-mediated radioresistance is dependent on p53. The colon cancer cell line DLD-1 harboring mutant p53 and Ki-Ras genes shows no difference in radioresistance compared to its counterpart DKO4, in which the mutant Ras has been deleted (Figure 5C), suggesting that Ras has no impact in cells lacking p53. However, it is important to note that DLD-1 and HCT116 cells do not share an identical genetic background. Nonetheless, this demonstrates an important role for oncogenic Ras mutations in the presence of wild-type p53 in conferring radioresistance in human tumors.

Taken together, these results provide strong evidence that activated H-Ras leads to increased basal Mdm2 protein expression, which is capable of attenuating p53 induction upon  $\gamma$ -irradiation and renders cells more resistant to the inhibitory effects of irradiation.

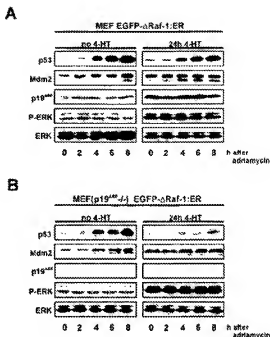
#### Attenuation of p53 Accumulation in Response to DNA-Damaging Reagents by Activated Raf Is Dependent on the p19<sup>ARF</sup> Status

p19<sup>ARF</sup> physically interacts with Mdm2 and consequently stabilizes p53 (Kamijo et al., 1998; Pomerantz et al., 1998; Stott et al., 1998). To investigate whether the presence of p19<sup>ARF</sup> can prevent Raf-induced Mdm2 from degrading p53, we established wild-type MEFs and p19<sup>ARF</sup> null MEFs expressing a conditionally active Raf allele (EGFP- $\Delta$ Raf1:ER) (Kamijo et al., 1997). MEFs and p19<sup>ARF</sup> null MEFs with induced and noninduced Raf kinase were treated with the DNA damaging reagent adriamycin (0.25  $\mu$ M). As shown in Figure 6, activation of Raf leads to a significant induction of Mdm2 protein after 24 hr in both cell lines. Concomitantly, p19<sup>ARF</sup> levels increase in 4-HT-treated wild-type MEFs expressing EGFP- $\Delta$ Raf1:ER (Figure 6A). While Raf-induced Mdm2 can attenuate p53 accumulation in response to DNA damage in p19<sup>ARF</sup> null MEFs, we detect no attenuation of p53 response in wild-type MEFs upon Raf activation (Figures 6A and 6B). These results provide evidence that p19<sup>ARF</sup> is capable of neutralizing the increased Mdm2 protein in response to activation of the Ras/Raf/MEK/MAP kinase pathway. If p19<sup>ARF</sup> is not expressed due to a targeted deletion of exon1 $\beta$  of the INK4A gene locus (Kamijo et al., 1997), induced Mdm2 protein is fully functional, can bind to p53, and promote its degradation.

#### Discussion

p53 is the major known regulator of its own inhibitor, Mdm2 (Barak et al., 1993; Wu et al., 1993). Mdm2-mediated degradation of p53 occurs in the cytoplasm, through a proteasome-dependent pathway (Haupt et al., 1997; Kubbutat et al., 1997). Hence, p53 and Mdm2 have been postulated to form an autoregulatory negative feedback loop. Regulation of *mdm2* expression by p53 is thought to keep p53 function under control, thereby preventing widespread p53-dependent apoptosis (Lane and Hall, 1997; Prives, 1998).

In this study, we demonstrate that Mdm2 expression is also modulated by the Ras/Raf/MEK/MAP kinase pathway through activation of Ets and AP-1 sites in the P2 promoter, upstream from the p53 responsive element and independent of its activity. Furthermore, Mdm2 induced by the Ras/Raf/MEK/MAP kinase pathway is



**Figure 6.** Attenuation of the p53 Accumulation by Raf-Induced Mdm2 in Response to DNA Damage Depends on the p19<sup>ARF</sup> Status. (A) Wild-type MEFs expressing a 4-Hydroxy-Tamoxifen (4-HT) inducible EGFP- $\Delta$ Raf1:ER construct were treated with 1  $\mu$ M 4-HT (right panel), or mock treated (left panel). After 24 hr, adriamycin (0.25  $\mu$ M) was added to the culture medium. Total protein was harvested at the indicated time points after addition of adriamycin and subjected to Western blot analysis. Lysates (+/- 4-HT) were run on the same gel. Exposure time is identical for each protein shown, allowing direct comparison of signal intensity. (B) p19<sup>ARF</sup> null MEFs expressing a 4-Hydroxy-Tamoxifen (4-HT) inducible EGFP- $\Delta$ Raf1:ER construct were treated with 1  $\mu$ M 4-HT (right panel), or mock treated (left panel). Twenty-four hours later, adriamycin (0.25  $\mu$ M) was added to the medium.

functionally active and leads to degradation of p53. This signaling pathway is intact in tumor cells expressing activated Ras as Mdm2 protein levels decrease dramatically after inhibiting MEK activity in these cells. Importantly, the effects of induced Mdm2 on p53 are regulated by p19<sup>ARF</sup>. Ras therefore acts on p53 through two competing pathways (Figure 7). Activation of the Ras/Raf/MEK/MAP kinase cascade results in elevated levels of Mdm2 protein. However, in normal cells, this pathway also induces the expression of p19<sup>ARF</sup> (Bates et al., 1998; Palmero et al., 1998), which inhibits Mdm2 activity (Tao and Levine, 1999; Sherr and Weber, 2000). Thus, in normal cells, levels of p53 are determined by a balance between opposing effects of the Ras/Raf/MEK/MAP kinase pathway. In MEFs, these opposing effects are equivalent, and Raf is ineffective at inducing p53, despite its effects in p19<sup>ARF</sup>. In different cell types, or even in MEFs growing under slightly different conditions, the balance of these opposing pathways is likely to be different. For example, in IMR90 human diploid fibroblasts, activated MEK leads to accumulation of p53, presumably because p14<sup>ARF</sup> exceeds Mdm2 induction (Lin et al., 1998). In Figure 7, DNA damaging agents are shown

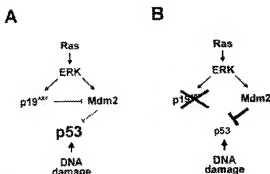


Figure 7. Model for the Regulation of p53 by the Ras/Raf/MEK/Map Kinase Pathway

(A) Control of p53 levels in wild-type cells. Activation of the Ras/Raf/MEK/Map kinase cascade leads to transcriptional induction of p19<sup>Arf</sup> by E2F-1 via the cyclin D/Cdk4/Rb pathway. Although Mdm2 protein is also induced, p19<sup>Arf</sup> is capable of keeping Mdm2 functionally inactive. p53 accumulates rapidly after DNA damage.

(B) Control of p53 levels in cells lacking p19<sup>Arf</sup>. Activated Ras/Raf/ERK kinase induces the transcription of *mdm2*. Because p19<sup>Arf</sup> is not expressed, the elevated Mdm2 protein is functionally active and attenuates p53 accumulation in response to DNA damage.

regulating p53 in a p19<sup>Arf</sup>-independent manner; however, it is possible that p19<sup>Arf</sup> affects the magnitude or duration of the p53 response to DNA under some physiological circumstances. In contrast to the Ras/Raf/MEK/Map kinase pathway, which activates opposing regulators of p53, E1A induces p19<sup>Arf</sup> but does not directly induce Mdm2 (de Stanchina et al., 1998). E1A is therefore a potent inducer of p53.

p53 function is lost by mutation of the p53 gene in about 60% of all human tumors (Hollstein et al., 1991; Levine et al., 1991). In many of the remaining tumors, p53 function is abrogated by overexpression of Mdm2 (Oliner et al., 1992), expression of HPV E6 (Schoffner et al., 1991), or by loss of p14<sup>Arf</sup> expression (Kamb et al., 1994; Cordon-Cardo, 1995; Haber, 1997). In about 30% of human tumors, Ras is activated by mutation (Bos, 1989). Although Ras mutation and p53 mutation seem to be independent events (Mitsudomi et al., 1992), Ras-induced Mdm2 might block p53 from inducing apoptosis or growth arrest in the early phase of tumor development allowing coexistence of Ras mutations and wild-type p53. For example, Ras mutations precede p53 mutations in the stepwise development of colon cancer (Kinzler and Vogelstein, 1990). Activation of Ras may suppress p53 during the early stages of tumor development. Moreover, some mutant forms of p53 may retain residual activity (Friedlander et al., 1996). Elevated Mdm2 protein levels induced by Ras activation may bind those mutated forms of p53 and abolish remaining p53 function. Consistent with this, we find that the Ras/Raf/MEK/Map kinase pathway suppresses expression of mutant p53 in DK4/4 colon cancer cells. In addition, p53-independent transformation properties of Mdm2 have been reported. Sarcomas harboring *mdm2* gene amplifications together with p53 mutations show worse prognosis, when compared to tumors with genetic alterations of p53 or *mdm2* alone (Cordon-Cardo et al., 1994). Therefore, Ras-induced Mdm2 might contribute to tumor progression in a p53-independent manner. Mouse models have

revealed that tumors in which p53 is lost through mutation or by loss of p19<sup>Arf</sup> are phenotypically similar, at least in the early stages of tumor development (Doherty et al., 1992; Kamijo et al., 1997). However, tumors that have lost p53 by direct mutation are genetically unstable relative to those that have lost p53 function through mutation of p19<sup>Arf</sup> (Kamijo et al., 1997). The latter tumors may retain residual p53 function that protects cells from genetic rearrangement. Indeed, p53 transcriptional activity can be measured easily in p14<sup>Arf</sup>-deficient tumor cells that retain low levels of p53 protein. In these tumors, activation of Ras may suppress residual p53 and allow tumor progression.

In addition to a role in tumor development, Ras regulation of *mdm2* and p53 may have important implications in cancer treatment. Tumors carrying constitutively active forms of Ras might be more resistant to treatment with ionizing radiation and chemotherapy. This idea is supported by previous observations that Ras-transformed NIH-3T3 cells are relatively resistant to radiation induced apoptosis (FitzGerald et al., 1985; Sklar, 1988). Chang and coworkers examined the radiation survival of various NIH-3T3 transformants representative of the various classes of oncogenes that may be involved in the pathway. In their study, Ras, Raf, Ets, and jun overexpressed in NIH-3T3 cells conferred a radiation resistant phenotype (Pirolo et al., 1993). Interestingly, chronic Ras transformation leads to an increase in AP-1 activity and upregulation of c-jun (Cook et al., 1998). As overexpression of transcription factors c-jun and Ets mediates radioresistance, a newly transcribed protein that functions as an inhibitor of radiation-induced apoptosis is likely to be involved. We provide evidence that Mdm2, a well characterized inhibitor of p53 activity, mediates the radioresistant phenotype conferred by oncogenic Ras. Our study demonstrates that constitutively active Ras induces *mdm2* transcription via activation of transcription factors binding to the AP-1 and Ets elements within the promoter. In the absence of p19<sup>Arf</sup>, elevated Mdm2 protein levels resulting from constitutively active Ras/Raf/MEK/Map signaling lead to attenuated and diminished p53 response and increased survival rates upon DNA damage. We propose that higher basal levels of Mdm2 protein in Ras-transformed cells prevents the accumulation of stable p53 protein and the subsequent induction of apoptosis or growth arrest in response to DNA damage.

Recently, we have suggested that high levels of Mdm2 activity explain why the E1B55k-deleted adenovirus d1520 (ONYX-015) replicates efficiently in many tumor cell lines that retain wild-type p53 (Ries et al., 2000). Oncogenic Ras, or other events that upregulate the MAP kinase pathway, could contribute to replication of d1520 in these cells by elevating levels of Mdm2 protein.

In vivo studies provide further evidence for the importance of Ras-controlled Mdm2 expression. Inhibition of oncogenic Ras activity in mouse models, through pharmacological intervention or genetic manipulation, leads to death by apoptosis (Lebowitz et al., 1997; Johnson et al., 1997; Heimbrook and Olf, 1998). This could be due, in part, to activation of p53 through reduced expression of Mdm2. To reevaluate the role of *mdm2* and p19<sup>Arf</sup> in Ras-induced tumor development in vivo, we are currently investigating the contribution of this pathway in a defined multistage mouse tumor model.

In conclusion, we have shown that Mdm2 is a transcriptional target of the Ras/Raf/MEK/Map kinase pathway, and that this activation is independent of p53. Ras

therefore regulates p53 through opposing pathways involving Mdm2 and its inhibitor p14<sup>ARF</sup>. In cancer cells lacking p14<sup>ARF</sup>, Ras suppresses p53 expression. This may have important implications in cancer development and therapy.

#### Experimental Procedures

##### Antibodies and Reagents

Rabbit polyclonal anti-phospho-ERK and anti-ERK antibodies were obtained from NERB, mouse monoclonal anti- $\beta$ -actin antibody was purchased from NID, and sheep polyclonal anti-p53 antibody Ab-7 was obtained from Calbiochem. Mouse monoclonal anti-Mdm2 antibody 2A10 was kindly provided by G. Zambetti (St. Jude Children's Hospital). The Dual Luciferase system and the MEK-inhibitor U0126 were from Promega.

##### Plasmids

Retroviral vectors (pBabe puro or pWZL3blast) expressing  $\Delta$ B-Raf-ER<sup>+</sup>, EGFP $\Delta$ Raf-1-ER proteins, or H-Ras(G12V) have been described elsewhere (Rodriguez-Viciana et al., 1997; Woods et al., 1997). Antisense riboprobe for RNase protection assay contains the murine *mdm2* cDNA fragment spanning from nt +264 to nt +3 (Barak et al., 1994).

##### Tissue Culture and Cell Lines

p19<sup>ARF</sup> null mouse embryo fibroblasts (Kamijo et al., 1997) and wild-type mouse embryo fibroblasts expressing the EGFP- $\Delta$ Raf-1-ER construct were used for experiments between passage 5 and 8 (Kamijo et al., 1997).

All cell lines were maintained in phenol-free DMEM H21 medium to prevent basal activation of the Raf/ER fusion protein. Asynchronous cell populations were treated with 10 nM 4-HT (for  $\Delta$ B-Raf-ER<sup>+</sup>) or 100 nM (for EGFP $\Delta$ Raf-1-ER) at 80% confluency for 24 hr or 48 hr, respectively. SW480 cells were maintained in Leibovitz medium supplemented with 10% heat-inactivated fetal bovine serum. To establish an inducible Mdm2 NIH-3T3 cell line, the commercially available pSwitch system (Invitrogen) was used. In radiation survival experiments, NIH-3T3 cells stably transfected with the pSwitch vector only treated with mitopretin ( $10^{-4}$  M) served as control.

##### Transfections and Luciferase Reporter Gene Assay

Reporter assays were performed as described previously using the Dual Luciferase system from Promega (Biedler et al., 2000). p53 activity was monitored using the PathDetect In Vivo Signal Transduction Pathway Cis-Reporting system (Stratagene, LaJolla, CA).

##### Northern Blot Analysis

Total cellular RNA was isolated using the RNeasy kit (Qiagen). Ten micrograms of total RNA per lane was subjected to a 0.9% agarose formaldehyde gel electrophoresis, transferred to Hybond N<sup>+</sup> membrane, and UV cross-linked (1200  $\mu$ J). Prehybridization, hybridization, and washing of blots were performed at 60°C in QuickHyb buffer (Amersham). The blots were stripped by boiling in 0.1% SDS, and then reprobed.

##### RNase Protection Assay

Ribonuclease protection assays were performed using the RPA kit (Ambion) and a cDNA riboprobe corresponding to exons 1–3 of the full-length *mdm2* transcript. *Mdm2* mRNA arising from the upstream promoter protects a fragment of 260 nt, whereas mRNA initiated from the internal promoter (P2) yields a fragment of 147 nt (Barak et al., 1994).

##### Western Blot Analysis

Cells were lysed in 2X Laemmli buffer and equal amounts of protein were separated by SDS-PAGE, transferred onto nitrocellulose membranes, and incubated with various antibodies. All immunoblots were visualized by enhanced chemiluminescence detection system (Amersham).

##### Site-Directed Mutagenesis

Misense mutations were engineered into the various *mdm2* promoter luciferase constructs by primer-mediated mutagenesis using the QuickChange mutagenesis kit (Stratagene). All mutations were verified by sequencing.

##### Electrophoretic Mobility Shift Assay

Double-stranded oligonucleotides spanning regions –129 bp to –91 bp (G51), and –171 bp to –136 bp (G52) of *mdm2* intron 1 (relative to the first nucleotide of exon 2) were labeled with [ $\gamma$ -<sup>32</sup>P]dCTP. Nuclear extracts of NIH-3T3 cells were prepared as described elsewhere (Andrews and Falter, 1991). EMSAs were carried out using 10  $\mu$ g nuclear extract as previously described (Moser et al., 1995).

##### Acknowledgments

We are grateful to G. Zambetti for the generous gift of Mdm2 antibody 2A10 and P. Rodriguez-Viciana for H-Ras(G12V) expression vector. The p19<sup>ARF</sup> null MEFs were generously provided by C. Sherr and M. Roussel. HCT116 p53<sup>-/-</sup> cells were a kind gift of F. Bunz and B. Vogelstein. We thank J. Ghysdael and M. Karin for expression vectors encoding c-ets-1, c-ets-2, and c-jun, F. Zindy for helpful advice with the mouse p53 antibody Ab-7, and M. Hult for DKO cells expressing  $\Delta$ Raf-ER. We are indebted to C. Brandts for critically reading the manuscript and to A. Balmain for exciting discussions. We thank P. Sabbatini, H. Jiang, O. Totsu, C. O'Shea, D. Stokoe, A. Bhat, and A. Mirza for fruitful discussions and all members of the McCormick lab for their support. S. Ries received a DAAD fellowship as part of the "Germanische Hochschulniederprogramm III von Bund und Landem". This study was supported by the David A. Wood Foundation.

Received May 30, 2000; revised September 5, 2000.

##### References

- Albanese, C., Johnson, J., Watanabe, G., Eklund, N., Vu, D., Arnold, A., and Pestell, R.G. (1995). Transforming p21ras mutants and c-Ets-2 activate the cyclin D1 promoter through distinguishable regions. *J. Biol. Chem.* 270, 23569–23577.
- Andrews, N.C., and Falter, D.V. (1991). A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res.* 19, 2199.
- Barak, Y., Juven, T., Hafner, R., and Oren, M. (1993). *mdm2* expression is induced by wild type p53 activity. *EMBO J.* 12, 461–468.
- Barak, Y., Gottlieb, E., Juven-Gershon, T., and Oren, M. (1994). Regulation of *mdm2* expression by p53: alternative promoter products transcripts with nonidentical translation potential. *Genes Dev.* 8, 1739–1749.
- Bates, S., Phillips, A.C., Clark, P.A., Stott, F., Peters, O., Ludwig, R.L., and Vousden, K.H. (1998). p14ARF links the tumour suppressors Rb and p53. *Nature* 395, 124–125.
- Biedler, C., Ries, S., Moser, M., Florio, M., Israel, M., McCormick, F., and Buehler, R. (2000). The basic helix-loop-helix transcription factors myogenin and Id2 mediate specific induction of the caveolin-3 gene expression during embryonic development. *J. Biol. Chem.* 275, 26245–26251.
- Boe, J.L. (1989). Ras oncogenes in human cancer: a review. *Cancer Res.* 49, 4682–4689.
- Bunz, F., Dutraux, A., Lengauer, C., Waldman, T., Zhou, S., Brown, J.P., Sedivy, J.M., Kinzler, K.W., and Vogelstein, B. (1998). Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* 282, 1467–1471.
- Cook, S.J., Aziz, N., and McMahon, M. (1999). The repertoire of fos and jun proteins expressed during the G1 phase of the cell cycle is determined by the duration of mitogen-activated protein kinase activation. *Mol. Cell Biol.* 19, 330–341.
- Cordon-Cardo, C., Latres, E., Drobnyak, M., Oliva, M.R., Pollack, D., Woodruff, J.M., Marochal, V., Chen, J., Brennan, M.F., and Levine, A.

- A.J. (1994). Molecular abnormalities of mdm2 and p53 genes in adult soft tissue sarcomas. *Cancer Res.* 54, 794-799.
- Cordon-Cardo, C. (1995). Mutations of cell cycle regulators, biological and clinical implications for human neoplasia. *Am. J. Pathol.* 147, 545-560.
- de Stanchina, E., McCormack, M.E., Zindy, F., Shieh, S.Y., Ferbeyre, G., Samuelson, A.V., Privet, C., Roussel, M.F., Sherr, C.J., and Lowe, S.W. (1998). E1a signaling to p53 involves the p19(ARF) tumor suppressor. *Genes Dev.* 12, 2436-2442.
- Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Jr., Butte, J.S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* 356, 215-221.
- Esteller, M., Tortola, S., Toyota, M., Capella, G., Peinado, M.A., Bayle, S.B., and Herman, J.G. (2000). Hypermethylation-associated inactivation of p14(ARF) is independent of p16(INK4a) methylation and p53 mutational status. *Cancer Res.* 60, 129-133.
- Fakhrazadeh, S.S., Trusko, S.P., and George, D.L. (1991). Tumorigenic potential associated with enhanced expression of a gene that is amplified in a mouse tumor cell line. *EMBO J.* 10, 1565-1569.
- Fambrough, D., McClure, K., Kazianka, A., and Lander, E.S. (1999). Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. *Cell* 97, 727-741.
- FitzGerald, T.J., Daugherty, C., Kane, K., Patenstein, L.A., McKenna, M., and Greenberger, J.S. (1995). Activated human H-ras oncogene enhances x-irradiation repair of mammalian cells in vitro less effectively at low dose rate. Implications for increased therapeutic ratio of low dose rate irradiation. *Am. J. Clin. Oncol.* 8, 517-522.
- Friedlander, P., Haupt, V., Privet, C., and Oren, M. (1996). A mutant p53 that discriminates between p53-responsive genes cannot induce apoptosis. *Mol. Cell. Biol.* 16, 4961-4971.
- Haber, D.A. (1997). Splicing into senescence: the curious case of p16 and p19ARF. *Cell* 91, 555-558.
- Haupt, V., Maya, R., Karz, A., and Oren, M. (1997). Mdm2 promotes the rapid degradation of p53. *Nature* 387, 296-299.
- Heimbrook, D.C., and Olt, A. (1999). Therapeutic intervention and signaling. *Curr. Opin. Cell Biol.* 10, 284-288.
- Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C.C. (1991). p53 mutations in human cancers. *Science* 253, 49-53.
- Honda, R., and Yasuda, M. (1999). Association of p19(ARF) with Mdm2 inhibits ubiquitin ligase activity of Mdm2 for tumor suppressor p53. *EMBO J.* 18, 22-27.
- Johnson, L., Greenbaum, D., Cichowski, K., Mercer, K., Murphy, E., Schmitt, E., Bronson, R.T., Linowski, H., Edelmann, W., Kaelin, P., and Jacks, T. (1997). K-ras is an essential gene in the mouse with partial functional overlap with N-ras. *Genes Dev.* 11, 2468-2481.
- Jones, S.N., Roe, A.E., Donehower, L.A., and Bradley, A. (1995). Rescue of embryonic lethality in Mdm2-deficient mice by absence of p53. *Nature* 378, 206-208.
- Juven, T., Barak, Y., Zauberman, A., George, D.L., and Oren, M. (1993). Wild type p53 can mediate sequence-specific transactivation of an internal promoter within the mdm2 gene. *Oncogene* 8, 3411-3416.
- Kern, A., Gris, N.A., Weaver-Felchhaus, J., Liu, G., Harshman, K., Tavtigian, S.V., Stockert, E., Day, R.S., 3rd, Johnson, B.E., and Skolnick, M.H. (1994). A cell cycle regulator potentially involved in genesis of many tumor types. *Science* 264, 438-440.
- Kamijo, T., Zindy, F., Roussel, M.F., Quelle, D.E., Downing, J.R., Ashman, R.A., Grossfeld, G., and Sherr, C.J. (1997). Tumor suppression at the mouse Hdk4 locus mediated by the alternative reading frame product p19ARF. *Cell* 91, 649-659.
- Kamijo, T., Weber, J.D., Zambetti, G., Zindy, F., Roussel, M.F., and Sherr, C.J. (1998). Functional and physical interactions of the ARF tumor suppressor with p53 and Mdm2. *Proc. Natl. Acad. Sci. USA* 95, 8292-8297.
- Katz, M.E., and McCormick, F. (1997). Signal transduction from multiple Ras effectors. *Curr. Opin. Genet. Dev.* 7, 75-79.
- Kinder, K.W., and Vogelstein, B. (1995). Lessons from hereditary colorectal cancer. *Cell* 87, 159-170.
- Kubbutat, M.H., Jones, S.N., and Vousden, K.H. (1997). Regulation of p53 stability by Mdm2. *Nature* 387, 299-303.
- Lane, D.P., and Hall, P.A. (1997). MDM2 arbiter of p53's destruction. *Trends Biochem. Sci.* 22, 372-374.
- Leibowitz, P.F., Sakamuro, D., and Prendergast, G.C. (1997). Farnesyl transferase inhibitors induce apoptosis of Ras-transformed cells denied substratum attachment. *Cancer Res.* 57, 708-713.
- Leng, P., Brown, D.R., Shivakumar, C.V., Deb, S., and Deb, S.P. (1995). N-terminal 130 amino acids of MDM2 are sufficient to inhibit p53-mediated transcriptional activation. *Oncogene* 10, 1275-1282.
- Leri, A., Liu, Y., Claudio, P.P., Kajitani, J., Wang, X., Wang, S., Kang, P., Mahdavi, A., and Anversa, P. (1999). Insulin-like growth factor-1 induces Mdm2 and down-regulates p53, attenuating the myocyte renin-angiotensin system and stretch-mediated apoptosis. *Am. J. Pathol.* 154, 567-580.
- Levine, A.J., Momand, J., and Finlay, C.A. (1991). The p53 tumour suppressor gene. *Nature* 351, 453-456.
- Lin, A.W., Barradas, M., Stone, J.C., van Aelst, L., Serrano, M., and Lowe, S.W. (1998). Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. *Genes Dev.* 12, 3008-3019.
- Mariash, M.S. (1995). Ras target proteins in eukaryotic cells. *FASEB J.* 9, 1311-1318.
- Martin, K., Trouche, D., Hagemeier, C., Sorensen, T.S., La Thangue, N.B., and Kouzarides, T. (1995). Stimulation of E2F1/DP1 transcriptional activity by MDM2 oncoprotein. *Nature* 375, 691-694.
- Mitsudomi, T., Steinberg, S.M., Nau, M.M., Carbone, D., D'Amico, D., Bodner, S., Oie, H.K., Linols, R.L., Mushinski, J.L., Minna, J.D., et al. (1995). p53 gene mutations in non-small-cell lung cancer cell lines and their correlation with the presence of ras mutations and clinical features. *Oncogene* 7, 171-180.
- Momand, J., Zambetti, G.P., Olson, D.C., George, D., and Levine, A.J. (1992). The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* 69, 1237-1245.
- Montes de Oca Luna, R., Wagner, D.S., and Lozano, G. (1995). Rescue of early embryonic lethality in mdm2-deficient mice by deletion of p53. *Nature* 378, 203-206.
- Moser, M., Imhof, A., Pscherer, A., Bauer, R., Amelgruber, W., Sinowatz, F., Hofstadter, P., Schüle, R., and Buitner, R. (1999). Cloning and characterization of a second AP-2 transcription factor: AP-2 beta. *Development* 121, 2779-2788.
- Oliner, J.D., Kinder, K.W., Meltzer, P.S., George, D.L., and Vogelstein, B. (1992). Amplification of a gene encoding a p53-associated protein in human sarcomas. *Nature* 358, 80-83.
- Oliner, J.D., Pretenet, J.A., Thaggling, S., Gyuris, J., Kinder, K.W., and Vogelstein, B. (1993). Oncoprotein MDM2 conceals the activation domain of tumor suppressor p53. *Nature* 362, 857-860.
- Palmer, L., Pantoja, C., and Serrano, M. (1998). p19ARF links the tumor suppressor p53 to Ras. *Nature* 395, 125-128.
- Peeper, D.S., Upton, T.J., Ludtke, M.H., Neuman, E., Zalcide, J., Bomards, R., DeCaprio, J.A., and Ewen, M.E. (1997). Ras signalling linked to the cell-cycle machinery by the retinoblastoma protein. *Nature* 386, 177-181.
- Picksley, S.M., and Lane, D.P. (1993). The p53-mdm2 autoregulatory feedback loop: a paradigm for the regulation of growth control by p53? *Bioessays* 15, 689-690.
- Pirolo, K.F., Tong, Y.A., Villegas, Z., Chen, Y., and Chang, E.H. (1993). Oncogene-transformed NIH 3T3 cells display radiation resistance levels indicative of a signal transduction pathway leading to the radiation-resistant phenotype. *Radiat. Res.* 135, 234-243.
- Pomerantz, J., Schreiber-Agus, N., Ueguchi, N.J., Silverman, A., Allard, L., Chin, L., Orlow, J., Chen, K., Lee, H.W., et al. (1998). The Ink4a tumor suppressor gene product, p19Arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53. *Cell* 92, 713-723.

- Prives, C. (1998). Signaling to p53: breaking the MDM2-p53 circuit. *Cell* 95, 5-6.
- Ries, S.J., Brandts, C.H., Chung, A., Biederer, C.H., Hann, B.C., Lipner, E.M., McCormick, F., and Korn, W.M. (2000). Loss of p14ARF facilitates replication of the adenovirus mutant dl1520 (ONYX-015) in tumor cells. *Nat. Med.*, in press.
- Rodriguez-Viciana, P., Wana, P.H., Khwaja, A., Martin, B.M., Pappin, D., Cox, P., Waterfield, M.D., Ridley, A., and Downward, J. (1997). Role of phosphoinositide 3-OH kinase in cell transformation and control of the actin cytoskeleton by Ras. *Cell* 89, 457-467.
- Samuels, M.L., Weber, M.J., Bishop, J.M., and McMahon, M. (1993). Conditional transformation of cells and rapid activation of the mitogen-activated protein kinase cascade by an estradiol-dependent human raf-1 protein kinase. *Mol. Cell. Biol.* 13, 6241-6252.
- Scheffner, M., Munger, K., Byrne, J.C., and Howley, P.M. (1991). The state of the p53 and retinoblastoma genes in human cervical carcinoma cell lines. *Proc. Natl. Acad. Sci. USA* 88, 5523-5527.
- Sharma, S., Schwärze-Waldhoff, I., Oberhuber, H., and Schaller, R. (1993). Functional interaction of wild-type and mutant p53 transcribed into human tumor cell lines carrying activated ras genes. *Cell Growth Differ.* 4, 861-869.
- Shaulian, E., Resnitzky, D., Shitman, O., Blandino, G., Amsterdam, A., Yayon, A., and Oren, M. (1997). Induction of Mdm2 and enhancement of cell survival by bFGF. *Oncogene* 15, 2717-2725.
- Sherr, C.J., and Weber, J.D. (2000). The ARF/p53 pathway. *Curr. Opin. Genet. Dev.* 10, 94-99.
- Shirasawa, S., Furuse, M., Yokoyama, N., and Sasazuki, T. (1993). Altered growth of human colon cancer cell lines disrupted at activated Ki-ras. *Science* 260, 85-88.
- Sklar, M.D. (1988). The ras oncogenes increase the intrinsic resistance of NIH 3T3 cells to ionizing radiation. *Science* 239, 645-647.
- Stokoe, C., Macdonald, S.G., Cudwallader, K., Symons, M., and Hancock, J.F. (1994). Activation of Raf as a result of recruitment to the plasma membrane. *Science* 264, 1463-1467.
- Stott, F.J., Bates, S., James, M.C., McConnell, B.B., Starborg, M., Brookes, S., Palmero, I., Ryan, K., Hara, E., Voudsen, K.H., and Peters, G. (1998). The alternative product from the human CDKN2A locus, p14(ARF), participates in a regulatory feedback loop with p53 and MDM2. *EMBO J.* 17, 5001-5014.
- Tao, W., and Levine, A.J. (1998). p19(ARF) stabilizes p53 by blocking nucleocytoplasmic shuttling of Mdm2. *Proc. Natl. Acad. Sci. USA* 95, 6937-6941.
- Woods, D., Parry, D., Chervinski, H., Bosch, E., Lees, E., and McMahon, M. (1997). Raf-induced proliferation or cell cycle arrest is determined by the level of Raf activity with arrest mediated by p21Cip1. *Mol. Cell. Biol.* 17, 5589-5591.
- Wu, X., Baylis, J.M., Olson, D., and Levine, A.J. (1993). The p53-Mdm2 autoregulatory feedback loop. *Genes Dev.* 7, 1126-1132.
- Xiao, Z.X., Chen, J., Levine, A.J., Modjtahedi, N., Xing, J., Sellers, W.R., and Livingston, D.M. (1995). Interaction between the retinoblastoma protein and the oncoprotein MDM2. *Nature* 375, 694-698.
- Zauberman, A., Flusberg, D., Haupt, Y., Barsky, Y., and Oren, M. (1995). A functional p53-responsive intronic promoter is contained within the human mdm2 gene. *Nucleic Acids Res.* 23, 2584-2592.
- Zindy, F., Eiselen, C.M., Randle, D.H., Kamijo, T., Cleveland, J.L., Sherr, C.J., and Roussel, M.F. (1998). Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. *Genes Dev.* 12, 2424-2435.



## Amplification of a gene encoding a p53-associated protein in human sarcomas

J. D. Oliner\*, K. W. Kinzler†, P. S. Meltzer‡, D. L. George§ & B. Vogelstein\*

\* The Johns Hopkins Oncology Center, 624 North Bond Street, Baltimore, Maryland 21205, USA

† Departments of Pediatrics and Radiation Oncology, University of Michigan Cancer Center, MSRB 11 C950B, 1350 West Medical Center Drive, Ann Arbor, Michigan 48109, USA

§ Department of Human Genetics, University of Pennsylvania, Clinical Research Building, Philadelphia, Pennsylvania 19104, USA

DESPITE extensive data linking mutations in the p53 gene to human tumorigenesis<sup>1</sup>, little is known about the cellular regulators and mediators of p53 function. MDM2 is a strong candidate for one such cellular protein; the *MDM2* gene was originally identified by virtue of its amplification in a spontaneously transformed derivative of mouse BALB/c cells<sup>2</sup> and the MDM2 protein subsequently shown to bind to p53 in rat cells transfected with p53 gene<sup>3,4</sup>. To determine whether MDM2 plays a role in human cancer, we have cloned the human MDM2 gene. Here we show that recombinant-derived human MDM2 protein binds human p53 *in vitro*, and we use MDM2 clones to localize the human MDM2 gene to chromosome 12q13-14. Because this chromosomal position appears to be altered in many sarcomas<sup>5-7</sup>, we looked for changes in human MDM2 in such cancers. The gene was amplified in over a third of 47 sarcomas, including common bone and soft tissue forms. These results are consistent with the hypothesis that MDM2 binds to p53, and that amplification of MDM2 in sarcomas leads to escape from p53-regulated growth control. This mechanism of tumorigenesis parallels that for virally-induced tumours<sup>8,9</sup>, in which viral oncogene products bind to and functionally inactivate p53.

To obtain human complementary DNA clones, a murine MDM2 cDNA probe was used to initiate cDNA walking in a human library (see legend to Fig. 1). Sequence analysis of 25 clones revealed several cDNA forms indicative of alternative splicing. The predominant human form is compared with its murine counterpart in Fig. 1. There was an open reading frame extending from the 5' end of the human cDNA sequence to nucleotide 1,784. Although this signal for translation initiation could not be unambiguously defined, the ATG at nucleotide 312 was considered the most likely position for several reasons. First, the sequence similarity between human and mouse MDM2 declined dramatically upstream of nucleotide 312. Second, an inverse polymerase chain reaction (PCR) was used in an attempt to acquire additional upstream cDNA sequence<sup>10</sup>. The 5' ends of the PCR-derived clones were very similar (within 12 base pairs) to the 5' ends of clones obtained from the cDNA library, indicating that the 5' end of the human MDM2 sequence shown in Fig. 1 may represent the 5' end of the transcript. Third, *in vitro* translation of the sequence shown in Fig. 1, beginning with the methionine encoded by the ATG at position 312, generated a protein similar in size to that observed in human cells (see below).

Comparison of the human and mouse MDM2 coding regions showed that they were 80.2% identical and shared a basic nuclear localization signal at codons 181 to 185 (ref. 11), several casein kinase II serine phosphorylation sites<sup>12</sup>, an acidic activation domain at codons 223 to 274 (ref. 13), and two metal-binding sites at codons 305 to 322 and 463 to 478, neither of which is highly related to known DNA-binding domains<sup>14</sup>.

To determine whether the human MDM2 protein could bind to human p53 protein *in vitro*, a human MDM2 expression vector was constructed from the cDNA clones (see legend to Fig. 2). RNA transcribed from this vector using T7 RNA polymerase was used to program a rabbit reticulocyte lysate. Although the predicted size of the protein generated from the construct was only 55.2K (*M<sub>r</sub>* 55,200, extending from the methionine at nucleotide 312 to nucleotide 1,784), protein translated *in vitro* migrated at ~90K. The MDM2 protein was not immunoprecipitated with antibodies against either the C-terminal or N-terminal regions of p53 (Fig. 2, lanes 2 and 3). But when *in vitro*-translated human p53 was mixed with the human MDM2 translation product, the anti-p53 antibodies precipitated MDM2 protein with p53 (Fig. 2, lanes 5 and 6). As a control, a protein of similar electrophoretic mobility (MCC<sup>15</sup>) was mixed with p53 and there was no coprecipitation (Fig. 2, lanes 8 and 9). When an *in vitro*-translated His-175 mutant form of p53 was mixed with human MDM2 protein, a similar coprecipitation of MDM2 and p53 proteins was also observed (data not shown).

Polyclonal rabbit antibodies were raised against an *Escherichia coli*-produced human MDM2-glutathione S-transferase fusion protein. The anti-MDM2 antibodies immunoprecipitated p53 when mixed with MDM2 protein (Fig. 2, lane 15) but failed to precipitate p53 alone (Fig. 2, lane 13).

To establish the chromosomal localization of human MDM2, somatic cell hybrids were screened, and a human-hamster hybrid containing only human chromosome 12 hybridized to the human MDM2 probe. Screening of hybrids containing portions of chromosome 12 (ref. 16) with the same probe narrowed the localization to chromosome 12q13-14. Because this region of chromosome 12 is often aberrant in human sarcomas<sup>5-7</sup>, southern blot analysis to evaluate whether MDM2 was genetically altered in such cancers. We found a striking amplification of MDM2 sequences in several of these tumours (see examples in Fig. 3, lanes 2, 3 and 5). Of 47 sarcomas analysed, 17 showed a 5-50-fold MDM2 amplification. These tumours included 7 of 13 liposarcomas, 7 of 22 malignant fibrous histiocytomas, 3 of 11 osteosarcomas, and 0 of 1 rhabdomyosarcoma. Five benign soft tissue tumours (lipomas) and seventy-four carcinomas (colorectal or gastric) were also analysed by Southern blotting and no amplification was seen.

We next determined whether this gene amplification was associated with increased expression. Because of RNA degradation in primary sarcomas, only the cell lines could be productively analysed by northern blotting. In the one available sarcoma cell line with MDM2 amplification, a single transcript of ~5.5 kilobases (kb) was observed (Fig. 4a, lane 1). The amount of this transcript was much higher than in a sarcoma cell line without amplification (Fig. 4a, lane 2) or in a carcinoma cell line (Fig. 4a, lane 3). When purified messenger RNA (rather than total RNA) from the sarcoma cell line was used for analysis, a human MDM2 transcript of 5.5 kb could also be observed (Fig. 4a, lane 4). Expression of the MDM2 RNA in the sarcoma with amplification was estimated to be at least 30-fold higher than that in the other lines examined. This was consistent with results from western blot analysis. A protein of *M<sub>r</sub>* ~90K was expressed at high levels in the sarcoma cell line with MDM2 amplification (Fig. 4b, lane 3), whereas no expression was evident in two sarcoma cell lines without amplification or in the carcinoma cell line (Fig. 4b, lanes 1, 2 and 4). Five primary sarcomas were also analysed by western blotting. Three primary sarcomas with amplification expressed the same sized protein as that in the sarcoma cell line (Fig. 4b, lanes 1-3), but no protein was observed in the two sarcomas without amplification (Fig. 4b, lanes 4 and 5).

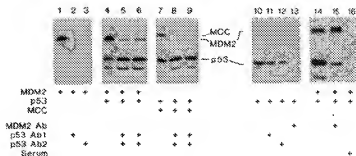
Our results demonstrate that human MDM2 binds to p53 *in vitro* and is genetically altered in a significant fraction of the most common sarcomas of soft tissue and bone<sup>17,18</sup>. It is important to note, however, that amplifications in human tumours

\* To which correspondence should be addressed.



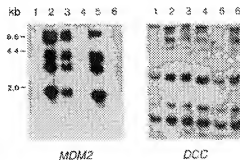
## LETTERS TO NATURE

**Fig. 2** Coprecipitation of human MDM2 and p53. In vitro translated MDM2, p53 and MCC proteins were mixed as indicated and incubated with p53 Ab1 (monoclonal antibody specific for the C terminus of p53), p53 Ab2 (monoclonal antibody specific for the N terminus of p53), MDM2 Ab (polyclonal rabbit anti-human MDM2 antibodies) or serum (preimmune serum obtained from the rabbit that produced the MDM2 antibody). Lanes 1, 4, 7, 10 and 14 contain aliquots of the protein mixtures used for immunoprecipitation. Bands running slightly faster than p53 are polypeptides produced from internal translation initiation sites. METHOS, A human MDM2 expression vector was constructed in pBluescript SK+ (Stratagene) from overlapping cDNA clones. The construct contained the sequence shown in Fig. 1 from nt 312 to 2,576. A 42-pb tickle vector ribosome entry sequence<sup>24</sup> was placed immediately upstream of this MDM2 sequence in order to obtain high expression. This construct, as well as p53 (ref. 26) and MCC<sup>13</sup> constructs in pBluescript SK+ were transcribed with T7 RNA polymerase and translated in a rabbit reticulocyte lysate (Promega) according to the manufacturer's instructions. Lysate (10 µl) containing the three proteins, alone or mixed in pairs, was incubated at 37 °C for 15 min. 1 µg (10 µl) of p53 Ab1 or Ab2 (Oncogene Science) or 5 µl of rabbit serum containing MDM2 antibody or preimmune rabbit serum, were added as indicated. 90 µl RPA buffer (10 mM Tris, pH 7.5, 1% sodium dodecylsulfate, 1% NP40, 150 mM NaCl, 0.1% SDS), SNANTE buffer<sup>25</sup>, or binding buffer<sup>26</sup> were then added and the mixtures allowed to incubate at 4 °C for 2 h. The three buffers produced similar results, although the

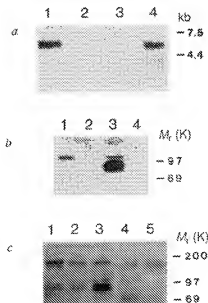


coprecipitation was less efficient in SNANTE buffer (containing 0.5 M NaCl; lanes 5 and 8) than in binding buffer (containing 0.1 M NaCl; lanes 6 and 9). Following addition of 2 mg protein A-Sepharose, the tubes were rotated end-over-end at 4 °C for 1 h. After pelleting and washing, immunoprecipitates were electrophoresed on SDS-polyacrylamide gels and the *in situ* gels autoradiographed in the presence of Enhance (New England Nuclear). Rabbits were immunized with a glutathione S-transferase (Pharmacia)/MDM2 fusion protein containing human MDM2 from the region corresponding to nt 350-816.

**Fig. 3** Amplification of the human MDM2 gene in sarcomas. DNA (5 µg) was digested with EcoRI, separated by agarose gel electrophoresis and transferred to nylon as described<sup>27</sup>. Filters were then hybridized with a human MDM2 cDNA fragment probe (nt 1-945; Fig. 1) or to a control probe that identifies fragments of similar size (pCC-1.65; ref. 28). Hybridization was as previously described<sup>28</sup>. DNA was derived from 5 primary sarcomas (lanes 1-4, 6) and one sarcoma cell line (lanes 5, 7). On longer exposures, the same sized MDM2 fragments were observed in lanes 1, 4 and 6. DNA fragment sizes are shown on the left in kb.



**Fig. 4** MDM2 expression. a, Northern blot analysis. RNA was separated by electrophoresis in a MOPS-formaldehyde gel and electrophoretically transferred to nylon filters. Transfer and hybridization were as described<sup>29</sup>. RNA was hybridized to the MDM2 fragment described in Fig. 3 legend. Total RNA (10 µg) was derived, respectively, from two sarcoma cell lines (OsA-G1, lane 2; and RC123, lane 2) and the carcinoma cell line (CaCo-2) used to make the cDNA library (lane 3). Lane 4 contains 2.0 µg polyadenylated CaCo-2 RNA. RNA sizes are shown on the right in kb. b, Western blot analysis of the sarcoma cell lines RC123 (lane 3), OsA-G1 (lane 3), HOS (lane 4), and the carcinoma cell line CaCo-2 (lane 2). c, Western blot analysis of primary sarcomas. Lanes 1 to 3 contain protein from sarcomas with MDM2 amplification, and lanes 4 and 5 contain protein from sarcomas without MDM2 amplification. Western blots using anti-purified MDM2 antibody were performed with 50 µg protein per lane as described<sup>24</sup>, except that the membranes were blocked in 10% non-fat dried milk and 10% goat serum, and secondary antibodies were coupled to horseradish peroxidase to allow chemiluminescent detection (Amersham ECL). MDM2 antibody was affinity-purified with a pATH-MDM2 fusion protein using methods described in ref. 31. Nonspecific reactive proteins of 75, 105 and 170R are seen in all lanes, irrespective of MDM2 amplification. MDM2 proteins of M<sub>r</sub> 90R were observed only in the MDM2-amplified tumours. Protein marker sizes are shown on the right.



(unpublished results with T. Tokino and D. Sidransky). The amplification of *MDM2* provides another provocative parallel between viral carcinogenesis and the naturally occurring genetic alterations underlying sporadic human cancer. □

Received 26 March accepted 5 May 1992

1. Melnicki, M., Saito, K., D. Vignani, R. A. Jones, C. C. Science **254**, 46-53 (1991).
2. Kohn, R., S. Tash, R. S. A. George, D. 1989, 1989-1991 (1991).
3. Jones, P. W. et al. *Dev. Cell* **1**, 571-580 (1991).
4. Nishida, K., Zarnegar, G. P., Olsen, D. G., George, D. L. & Levine, A. J. *Cell* **64**, 119-128 (1991).
5. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
6. Tash, D. G. et al. *Cell* **64**, 119-128 (1991).
7. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
8. Jones, P. W. et al. *Cell* **64**, 119-128 (1991).
9. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
10. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
11. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
12. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
13. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
14. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
15. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
16. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
17. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
18. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
19. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
20. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
21. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
22. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
23. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
24. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
25. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
26. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
27. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
28. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
29. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
30. Nishida, K. et al. *Cell* **64**, 119-128 (1991).
31. Nishida, K. et al. *Cell* **64**, 119-128 (1991).

ACKNOWLEDGEMENTS We thank A. Levine for the information about MDM2 p53 interaction that stimulated this work. M. L. Low for the p53 cDNA, and M. Levine for the p53 cDNA. We also thank Dr. S. Tash for the p53 cDNA. This work was supported by the National Institutes of Health.

## Wild-type p53 activates transcription *in vitro*

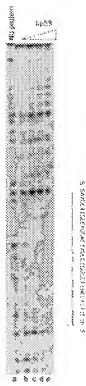
George Farmer, Jill Bargonetti, Hua Zhu, Paula Friedman, Ron Prywes & Carol Prives

Department of Biological Sciences, Columbia University, New York 10027, USA

THE p53 protein is a important determinant in human cancer and regulates the growth of cells in culture<sup>1-3</sup>. It is known to be a sequence-specific DNA-binding protein<sup>4,5</sup> with a powerful activation domain<sup>6-8</sup>, but it has not been established whether it regulates transcription directly. Here we show that intact purified wild-type human and murine p53 proteins strongly activate transcription *in vitro*. This activation depends on the ability of p53 to bind to a template bearing a p53-binding sequence. By contrast, tumour-derived mutant p53 proteins cannot activate transcription from the template at all, and when complexed to wild-type p53, these mutants block transcriptional activation by the wild-type protein. Moreover, the simian virus 40 large T antigen inhibits wild-type p53 from activating transcription. Our results support a model in which p53 directly activates transcription but this activity can be inhibited by mutant p53 and SV40 large T antigen through interaction with wild-type p53.

A DNA-binding immunassay has been used to screen human genomic clones and show that p53 binds specifically to a region upstream of the transcription start site for the human ribosomal gene cluster (RGC)<sup>9</sup>. We have confirmed and extended this observation by DNase I footprinting and shown that addition of immunoprecipitated p53 to a DNA fragment containing the RGC

FIG. 1 The p53 protein binds specifically to a site in the human ribosomal gene cluster. DNA binding was assayed in 50-µl volumes containing 40 mM creatine phosphate, pH 7.7, 4 mM ATP, 2 mM MgCl<sub>2</sub>, bovine serum albumin (2 mg ml<sup>-1</sup>), 0.5 mM dithiothreitol, 10 ng carrier plasmid (pT153) and 10 fmol of 5'-<sup>32</sup>P-labelled DNA fragment containing the ribosomal gene cluster (RGC) p53-binding site<sup>9</sup> and either no protein (lane a) or increasing amounts of wild-type human p53 in increments of 15 ng up to 50 ng (lanes b-e). DNase I treatment of mixtures and immunoprecipitation of samples for electrophoresis on 8% polyacrylamide urea gels has been described<sup>9</sup>. Wild-type p53 was immunoprecipitated from B27 cells expressing a recombinant oculo virus, pEV5bwt, using the monoclonal antibody Pab421 crosslinked to Sepharose A<sup>10</sup>.



site leads to strong and specific protection of only the RGC region (Fig. 1). All tumour-derived mutant p53 proteins tested failed to protect this sequence (J.B. *et al.*, manuscript in preparation).

To determine whether p53 can activate transcription *in vitro*, we used as templates the plasmids *fos*wt and *fos*lmt, which contain the human RGC p53 DNA-binding fragment or a mutated RGC fragment respectively (Fig. 2a). Three partially purified fractions from HeLa cell nuclear extracts were used as a source of transcription factors<sup>11</sup>. RNA products were analysed by S1 nuclease digestion using specific probes for each construct. Increasing amounts of p53 stimulated transcription from *fos*wt (compare lanes 1-4 with lanes 5-8). These reaction mixtures also included a construct containing an abridged adenovirus major late promoter (pMLS, ref. 10) whose transcription was not significantly affected by p53.

The p53 protein activated transcription from another promoter as well (Fig. 2b). Plasmids containing either one or sixteen copies of the RGC site, or one mutant RGC site, inserted adjacent to the polyoma virus early promoter to create Py1wt, Py16wt and Py1mt, respectively, were used as templates in transcription reactions. We found that p53 activated transcription of constructs containing the wild-type RGC (lanes 1-9) but not the mutant RGC (lanes 10-12). Diagrams of the templates and the test probe used in these experiments are shown in Fig. 2c with the expected S1 nuclease products.

The high incidence of p53 gene mutations in cancer patients suggests that alteration of the normal function of p53 is an important part of the oncogenic process. Therefore it was of interest to examine whether tumour-derived mutant p53 proteins activate transcription. The mutant p53 proteins we chose are defective in both nonspecific and specific DNA binding<sup>12,13</sup>, so providing an opportunity to confirm that p53 must bind DNA to activate transcription. We compared the ability of wild-type and two tumour-derived mutant p53 proteins with mutations at either amino acid 175 (His 175) or at amino acid 273 (His 273)



## Genomic organisation of the human *MDM2* oncogene and relationship to its alternatively spliced mRNAs<sup>☆</sup>

Huiling Liang<sup>a,1</sup>, Helen Atkins<sup>a</sup>, Rana Abdel-Fattah<sup>a</sup>, Stephen N. Jones<sup>b</sup>, John Lunec<sup>a,\*</sup>

<sup>a</sup> Cancer Research Unit, Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, UK

<sup>b</sup> University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, USA

Received 26 January 2004; received in revised form 26 April 2004; accepted 17 May 2004

Available online 21 July 2004

Received by A.J. van Wijnen

### Abstract

The *MDM2* proto-oncogene, which encodes a protein that binds to the p53 tumour suppressor, has been found amplified and overexpressed in a range of human tumours. Although the human *MDM2* cDNA sequence has been reported, the genomic organisation of the human gene has not been documented. We have previously reported the detection of five alternative internally deleted *MDM2* transcripts in human tumours and suggested these may represent alternatively spliced forms. Here we demonstrate two novel *MDM2* transcripts with internal deletions, using RT-PCR followed by sequencing. To definitively ascribe these variant transcript forms to alternative splicing, and to explore associated mechanisms, we have determined the intron–exon organisation of the human genomic sequence. The human *MDM2* gene spans approximately 33 kb and is divided into 12 exons. Exon sizes range from 50 to ≥ 1161 bp and intron sizes vary from 121 to ~ 7000 bp. The positions of intron–exon boundaries are compared with the deletion junctions of the multiple-sized transcripts and discussed in relation to alternative splicing mechanism.

© 2004 Elsevier B.V. All rights reserved.

**Keywords:** p53; genomic mapping; long range PCR

### 1. Introduction

The *mdm2* proto-oncogene was initially identified as an amplified gene from a mouse double minute chromosome present in a spontaneously transformed Balb/C 3T3 cell line, 3T3DM (Haines et al., 1994; Sigalas et al., 1996; Steinman et al., 2004). The causal role of this gene in tumorigenesis was originally established by transfection studies using genomic DNA sequences. In these studies,

experimental overexpression of *mdm2* resulted in the immortalisation of primary rat embryo fibroblasts and induced a fully transformed phenotype in the cells when cotransfected with an activated *ras* gene (Finlay, 1993). The human homologue of the *MDM2* gene has been found to be amplified in over 30% of human sarcomas (Oliner et al., 1992; Leach et al., 1993), which consequently results in high levels of the *MDM2* gene product. In addition, *MDM2* overexpression can also occur through enhanced transcription and translation (Bueso-Ramos et al., 1993; Landers et al., 1997; Momand et al., 1998).

The human *MDM2* gene has been localised to chromosome 12q13–14. Although the human *MDM2* cDNA sequence has been previously reported (Oliner et al., 1992), little is known about its genomic organisation. The *MDM2* protein is composed of 491 amino acids and contains a p53 binding domain (codons 19–102), a putative nuclear localisation signal (codons 181–185), an acidic domain (codons 223–274), a central zinc-finger motif (codons 305–332) and a ring-finger motif towards the C-terminal end of the protein (codons 438–478) (Boddy et al., 1994).

**Abbreviations:** *MDM2*, mouse double minute 2 gene; RT-PCR, reverse transcription polymerase chain reaction; cDNA, complementary DNA; kb, kilobase; bp, base pair; TE, triethoxysilyltrimethylammonium ethylene diamine tetra acetate.

<sup>☆</sup> Database accession numbers: AF144014–AF144033 and AF201370–AF201371.

<sup>\*</sup> Corresponding author. Northern Institute for Cancer Research Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, UK. Tel.: +44-191-246-4420; fax: +44-191-246-4391.

E-mail address: [john.lunec@newcastle.ac.uk](mailto:john.lunec@newcastle.ac.uk) (J. Lunec).

<sup>1</sup> Present address: University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, USA.

MDM2 appears to be a pluripotential oncoprotein, exerting its transforming properties through several alternative mechanisms, of which the most extensively studied has been the negative regulation of p53 function. MDM2 blocks p53 transcriptional function by binding to p53 (Momand et al., 1992; Oliner et al., 1993). The binding of MDM2 to p53 also results in the rapid degradation of p53 (Haupt et al., 1997; Kubbutat et al., 1997). In addition, MDM2 has been reported to have p53-independent tumorigenic properties. This includes the ability to interact with and inactivate the pRb tumour suppressor protein (Xiao et al., 1995) and to bind to and activate the E2F1 transcription factor (Martin et al., 1995). Furthermore, two independent transgenic studies have shown *MDM2* to have tumorigenic properties in p53 null mice (Lundgren et al., 1997; Jones et al., 1998).

One of the distinctive properties of *MDM2* is the possession of an extremely complex expression pattern. Its multiple-sized transcripts and proteins have been found in tumour samples and cell lines by a number of groups (Haines et al., 1994; Sigalas et al., 1996; Bartel et al., 2002). In our previous studies, five alternatively sized transcripts of the human *MDM2* were found in human ovarian tumour, bladder tumour and leukaemic cell samples (Sigalas et al., 1996). The expression of the alternatively sized forms was found to be more frequent in tumours of advanced stage and high histological grade, and they also retained their ability to transform NIH3T3 cells. Here, we present data demonstrating two further *MDM2* transcript forms with internal sequence deletions in human tumour tissue. We hypothesised that these transcripts are generated by alternative splicing. To test this hypothesis and to explore the associated mechanisms, we have investigated the genomic structure and organisation of the human *MDM2* gene. This gene is ~33 kb in length and comprises at least 12 exons. The sizes of exons vary from 50 to ≥1161 bp, and introns range in size from 121 to ~7000 bp. The position of intron–exon boundaries is compared with the sequences of the *MDM2* variant transcripts and discussed in relation to alternative splicing mechanisms.

## 2. Materials and methods

### 2.1. Nested RT-PCR

Total RNA was extracted from human bladder tumour and normal bladder tissues. Nested RT-PCR was carried out as previously described (Sigalas et al., 1996).

### 2.2. Genomic DNA extraction

Genomic DNA was prepared from frozen normal human placental tissue by digestion with proteinase K and phenol–chloroform extraction. The DNA was precipitated with a half volume of 7.5M ammonium acetate and one volume of isopropanol, washed in 70% ethanol and resuspended in 1 × TE buffer (10 mM Tris, 0.1 mM EDTA pH 7.5).

### 2.3. Long-range PCR

Primers (Table 1) were designed from the published *MDM2* cDNA sequence (Oliner et al., 1992). Each primer pair was designed to cross the deletion junctions of multiple-sized transcripts (Sigalas et al., 1996) or according to the predicted exon/intron boundaries by referring to the mouse *mdm2* gene structure (Jones et al., 1996; Montes de Oca Luna et al., 1996). A long-range PCR protocol was carried out, with the above human genomic DNA as a template, using an XL PCR Kit (Perkin Elmer, Part No. N808-192). For comparison, PCR was also carried out on normal human placental cDNA. The reaction contained 1 × reaction buffer, 0.8 mM dNTP, 1.1 mM Mg(OAc)<sub>2</sub> and 4 units of *rTth* DNA polymerase, 40 pmol of each primer and 100 ng of genomic DNA or 20 µl cDNA in a total volume of 100 µl. The long-range PCR was performed using a thermal cycler (Perkin-Elmer Model 480) as follows: 94 °C for 2 min; cycles 1–16 at 94 °C for 30 s, 58–62 °C for 10 min; cycles 17–28 at 94 °C for 30 s, 58–62 °C for 10 min and 15 s of increment

Table 1

The sequences of primer pairs for amplification of genomic nucleotides of the human *MDM2* gene

Primer name	Forward primer (5' → 3')	cDNA position (nt)	Reverse primer (3' → 5')	Position at cDNA (nt)	Annealing temperature (°C)
Pr1	ccctggtgcccctggtg	30–49	tgtctcgaagctggaactctg	382–361	60
Pr2	gfgcctctctctctctctg	314–333	ctacagactctctctctctctgaag	435–410	60
Pr3	gccttggagctctctctctg	410–435	ctacatctgtgtgtctctctctc	536–513	57
Pr4	ttttcttggccctctctctg	474–497	atctctgtgtgtgtctctctctc	654–632	57
Pr5	ctctctctctctctctctgag	614–636	tgtctctctctctctctctctg	716–695	60
Pr6	actcagctctctctctgag	661–683	tgtctctctctctctctctctc	815–793	60
Pr7	ttctctctctctctctctctg	724–746	atgtctctctctctctctctctc	969–949	60
Pr8	tgagactctctctctctg	893–912	caactctctctctctctctctg	1059–1036	57
Pr9	tctgagctctctctctctg	980–1002	agctctctctctctctctctctc	1211–1188	57
Pr10	gtgtctctctctctctctctg	1168–1191	ctctctctctctctctctctg	1277–1254	60
Pr11	ctttgtctctctctctctctg	1213–1235	cgtctctctctctctctctctc	2372–2351	57

\* Sequence of *MDM2* cDNA clones (Oliner et al., 1992).

per cycle, with a final extension of 72 °C for 10 min after the last cycle.

PCR products were separated by electrophoresis on a 1% low melting temperature agarose gel (NuSieve GTG, Flowgen) and visualised by ethidium bromide staining with UV transillumination. DNA bands amplified from genomic DNA were excised and purified with a QIAquick Gel Extraction kit (Qiagen), by following the procedure recommended by the company.

#### 2.4. Cloning PCR products

Purified PCR products were subcloned directly into the pGEM-T Easy vector (Promega) following the protocol recommended by the company. Ligation products were transformed into *Escherichia coli* JM109 and clones containing the desired inserts were identified by PCR screening. Plasmids were prepared by using the Wizard Plus SV Miniprep system (Promega).

#### 2.5. Sequencing

Sequencing was carried out manually by using PCR product directly as a template or automatically by using plasmid PCR product clone as a template. Manual sequencing was performed using the Sequenase Version 2.0 DNA sequencing system (Amersham, Product No.70770). The automated sequencing was carried out in the central core facility at the University of Newcastle upon Tyne Medical Faculty. The sequences were aligned with the published *MDM2* cDNA sequence, using the DNASTAR sequence analysis software package.

### 3. Results

#### 3.1. The detection of multiple-sized *MDM2* transcripts in human tumours

In our previous studies, we have found five alternative-sized *MDM2* transcripts (*MDM2-a*, *-b*, *-c*, *-d* and *-e*) (Sigalas et al., 1996). Our present investigation of the *MDM2* transcriptional pattern in human bladder tumour samples, but not in normal bladder tissue (data not shown), using RT-PCR has revealed two further transcripts, sized 813 and 707 bp, which we have designated *MDM2-a1* and *-g* (Fig. 1). Sequencing shows that these two transcripts have internal sequence deletions: *MDM2-a1* lacks nucleotides from codons 28 to 222 and codons 275 to 300; *MDM2-g* misses nucleotide codons from 28 to 97<sup>2/3</sup> and 114<sup>2/3</sup> to 300. Fig. 2 shows the structure of these transcripts in relation to the full-length *MDM2* cDNA sequence and previously described variant transcripts (see GenBank accession numbers AF201370 and AF201371 for details of the sequences).

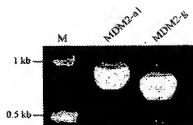


Fig. 1. *MDM2* transcripts were amplified by RT-PCR with nested primers that flank the *MDM2* coding region, as previously described (Sigalas et al., 1996; Matsumoto et al., 1998). M: molecular weight marker.

#### 3.2. *MDM2* genomic structure and organization

We have used long-range PCR amplification, followed by cloning and sequencing to investigate the organisation of the human *MDM2* gene and in particular to define intron–exon boundaries and flanking intronic sequences. Eleven DNA fragments were amplified from genomic DNA with the primer pairs shown in Table 1, which match to the known *MDM2* cDNA sequences. Comparison of the sequences of these PCR products with the published sequences of *MDM2* cDNA clones reveals that the Prii primer pair spans introns 1 and 2; while primer pairs Pri2–10 cover one intron per primer pair. However, the region flanked by primer pair Pri11, covering from the 1235th to 2351st nucleotide of the *MDM2* cDNA clone sequence (Oliner et al., 1992), was found not to contain any intronic sequence. Sequence analysis indicates that *MDM2* spans approximately 33 kb of genomic DNA and is separated by 11 introns. Exons range in size from 50 to  $\geq 1161$  bp. The size of the introns varies from 121 to  $\sim 7000$  bp (Table 2). Exon–intron boundary sequences of the 5' and 3' splice sites follow the "GT and AG" rule (Table 3; see GenBank accession numbers AF144014 AF144033 for additional intronic sequence data). Table 4 shows the 3' ends of the intronic sequences adjacent to the intron–exon boundaries, including branch sites and polypyrimidine tracts. The sequences of branch sites have a good match with the consensus sequence YURAY (Y: pyrimidine, R: purine, A\*: branch point residue). The distances between the branch points and the 3' splice sites vary from 18 to 111 bp. The C/T content in the polypyrimidine tracts ranges from 53% to 90%.

Comparison of the structure and organisation of the human *MDM2* gene described here with that published for the mouse gene (Jones et al., 1996) indicates that the number of the exons and introns is the same. The size of the coding exons is similar (Table 2). However, the sizes of the noncoding regions, including exons 1, 2 and 12 and the introns, differ substantially, with the exception of introns 1 and 3.

#### 3.3. Analysis of alternatively spliced variants of *MDM2* mRNA

We have detected seven *MDM2* transcript variants (*MDM2-a*, *-a1*, *-b*, *-c*, *-d*, *-e* and *-g*) previously and





Table 3  
Sequence of the human *MDM2* exon/intron boundaries

Exon no.	Sequence of boundary	Exon no.
1	GGAGCAGgtgctggc--intron 1--tttccacagCTGTGTT	2
2	GATCCAGgtaegcac--intron 2--ccttgtagGCAAAATG 1	3
3	GACCTCTgttagtat--intron 3--tcttatagGTTAGAC 27 28	4
4	GAAGGAGstaagctg--intron 4--tatttcagGTCTTT 52 53	5
5	AGGACAGgtaattct--intron 5--tctacaagGAAATA 96 98	6
6	CAGCAGgtaagtta--intron 6--tctctcagGAATCATC 113 115	7
7	TCAAAGGtaatacta--intron 7--atgcttagGACCTTG 136 137	8
8	GAGACAGgtatatat--intron 8--statccacAAGGAA 167 169	9
9	GAATCAGgtaagtgt--intron 9--tggttttagGAATCTG 222 223	10
10	TGATCAGgtatatat--intron 10--tctattagGATATC 274 275	11
11	CTTAGCTgtaagtat--intron 11--cattgaagGACTATT 300 301	12

Exon sequences are shown as uppercase letters. Intron sequences are denoted by lowercase letters. Underlined nucleotides encode amino acids (numbered underneath).

exon 6, has a shifted reading frame after codon 28. The *MDM2-e* variant, which also involves a deletion junction with an interrupted codon sequence, has a shift in the reading frame after codon 484.

#### 4. Discussion

In our previous studies and the results presented here, we detected multiple-sized *MDM2* transcripts in human

Table 4  
Branch sites and polypyrimidine tracts of the human *MDM2* gene

[illegible]

Branch site sequences are indicated with underlined letters. †: boundary between intron and exon.

<sup>a</sup> The percentage of T and C content in polypyrimidine tracts.

tumour tissues, but not in normal tissues. Detection of alternative spliced forms of *MDM2* mRNAs varied, and appeared to be relatively abundant in most samples. These variants encoded protein products in vitro and were found to transform NIH3T3 cells and to be associated with high-grade and late-stage human cancer (Sigalas et al., 1996 and unpublished data). Our data were supported by the observations of other groups, who reported that alternatively spliced *mdm2* transcripts promoted tumour formation in mouse model (Fridman et al., 2003; Steinman et al., 2004). These collective observations suggest that alternatively spliced forms of *MDM2*, encoding alternative proteins with differing functional capabilities, may play an important role in tumour development. The human *MDM2* genomic map presented here enable us to relate the genomic structure and organisation of *MDM2* to the appearance of variant transcript forms and provide a basis for considering alternative splicing mechanisms.

Pro-mRNA splicing involves precise excision of intron sequences and the ligation of exon sequences. In cases of alternative splicing, the excision may occur at cryptic splice sites; exons may be skipped and introns may be retained. The organisation of the exon–intron boundaries of the human *MDM2* gene indicates that the *MDM2-a*, *-a1*, *-b*, *-c* and *-g* variant transcript forms (Fig. 2) result from multiple entire exon skipping, because the internal deletion junctions correspond exactly to the location of exon–intron boundaries. However, the deletion junctions of *MDM2-d* and *-e* forms do not correspond to the boundaries between exons and introns, and there are no consensus splicing sequences surrounding them to indicate the possible use of cryptic splice sites. This suggests that they may have resulted from an unusual and possibly aberrant splicing mechanism.

The regulation of alternative splicing involves both *cis* elements and *trans*-acting factors. The *cis* elements include the 5' and 3' splice sites, a branch site and a polypyrimidine tract between the branch point and the 3' splice site. It has been demonstrated that a short distance between the branch point and the 3' splice site and high C/T content at polypyrimidine tracts give rise to high efficiency of splicing in mammals (Heliman et al., 1988; Libri et al., 1989). Our data show that all the 5' and 3' splice sites of the human *MDM2* gene obey the "GT" and "AG" rule. The branch sites also have a good match with the consensus sequence. However, the distances between branch points and 3' sites, and the percentage C/T content vary between introns. The exons most frequently retained in the splice variants have shorter distances between their upstream branch points and 3' splice sites and/or a higher percentage C/T content in their upstream polypyrimidine tracts, compared with those exons commonly excluded (Table 2). It suggests that the short distance between the branch point to the 3' site and the polypyrimidine tract with high C/T content confer

high splicing efficiency to their adjacent splicing sites, which is consistent with the observation reported by Smith et al. (1989), Goux-Pelletier et al. (1990), Mueller et al. (1997). However, exon 7 was spliced out with a high frequency although there is only a short distance between its branch point and 3' splice site and there is also a high C/T content in its upstream polypyrimidine tract. The explanation may be that it is spliced out along with upstream exons. The order of intron removal is governed by preferential binding of splice factors rather than in a sequential numerical order (Lewin, 1994). It may be possible that exon 7 processes splicing before exon 6 does, and the 5' end of exon 7 may in some circumstances be ligated preferentially to the 3' end of exon 6. If exon 6 has low splice strength, then in the process of ligation with its upstream exon, it may be out-competed and skipped out together with exon 7.

Currently, we are not clear why the alternative spliced transcripts appear preferentially in tumour samples, especially in advanced stage and high grade, but not in normal tissues (Sigalas et al., 1996; Bartel et al., 2002). Although variant *MDM2* spliced transcripts have been reported in normal tissues in one study (Bartel et al., 2004), we failed to detect these isoforms in noncancerous tissues. It has been proposed that a mRNA surveillance system exists in cells, which protects them from errors of transcription, mRNA processing, or mRNA transport (Pulak and Anderson, 1993). Mistakes are not uncommon in splicing of RNA from complex genes. Exons can occasionally be skipped (Nigro et al., 1991). In the normal situation, the surveillance system would probably degrade most mRNA with splicing errors as they are transported to the cytoplasm. We speculate that in cancer cells, this system may not function correctly and, consequently, the splice variants may escape degradation. It is also possible that there are mutations in the intron region that cause alternative splicing and the presence of the variants contribute to the cancer. It would be of interest to test this hypothesis by investigating intron nucleotide sequences in tumour samples that show expression of alternatively spliced forms. However, other models can be envisaged; for instance, we cannot rule out the possibility that *trans*-acting factors are involved in the alternative RNA processing by blocking some splice sites and/or enhancing other splice sites.

In conclusion, we have detected two novel *MDM2* alternatively spliced transcripts and have also defined the structure and organisation of the human *MDM2* gene. In addition we have related this information to potential mechanisms by which alternatively sized *MDM2* transcripts are generated. As the alternatively spliced *MDM2* mRNAs have been shown to possess oncogenic potential and to correlate with advanced malignancies to tumour progress (Haimes et al., 1994; Sigalas et al., 1996; Steinman et al., in press), our data may assist in clinical diagnosis of sarcomas displaying *MDM2* amplification and alternative splicing of *MDM2* transcripts.

## Acknowledgements

We thank Ratchada Suaeayon for her assistance with the work presented in this paper.

## References

- Bartel, F., Tauber, H., Harris, L., 2002. Alternative and aberrant splicing of MDM2 mRNA in human cancer. *Cancer Cell* 2, 9–15.
- Bartel, F., Finkert, D., Fiedler, W., Kappler, M., Wurl, P., Schmidt, H., Tauber, H., 2004. Expression of alternatively and aberrantly spliced transcripts of the MDM2 mRNA is not tumor-specific. *Int. J. Oncol.* 24, 143–151.
- Boddy, M.N., Freemont, P.S., Borden, K.L., 1994. The p53-associated protein MDM2 contains a newly characterized zinc-binding domain called the RING finger. *Trends Biochem. Sci.* 19, 198–199.
- Bueso-Ramos, C., Yang, Y., Delcon, E., McCown, P., Stass, S., Altshuler, M., 1993. The human MDM-2 oncogene is overexpressed in leukemias. *Blood* 82, 2617–2623.
- Finlay, C.A., 1993. The mdm-2 oncogene can overcome wild-type p53 suppression of transformed cell growth. *Mol. Cell Biol.* 13, 301–306.
- Fridman, J., Hernandez, E., Hermann, M., Stachura, E., Corfow-Dardo, S., Lowe, S., 2003. Tumor promotion by Mdm2 splice variants unable to bind p53. *Cancer Res.* 63, 5702–5706.
- Gnax-Pelletier, M., Libri, D., d'Aubenton-Carieu, Y., Fuzman, M., Brody, E., Marie, J., 1990. In vitro splicing of mutually exclusive exons from the chicken B tropomyosin gene: the role of the branch point location and very long pyrimidine stretch. *EMBO J.* 9, 241–249.
- Haines, D.S., Landers, J.B., Engle, L.J., George, D.L., 1994. Physical and functional interaction between wild-type p53 and mdm2 proteins. *Mol. Cell Biol.* 14, 1171–1178.
- Haupt, Y., Maya, R., Kazan, A., Oren, M., 1997. Mdm2 promotes the rapid degradation of p53. *Nature* 387, 296–299.
- Hellman, D.M., Ricci, W.M., Finn, L.A., 1988. Alternative splicing of tropomyosin pre-mRNAs in vitro and in vivo. *Genes Dev.* 2 (12A), 1827–1838.
- Jones, S.N., Anasari-Lari, M.A., Hancock, A.R., Jones, W.J., Gibbs, R.A., Donohower, L.A., Bradley, A., 1996. Genomic organization of the mouse double minute 2 gene. *Genes* 175, 209–213.
- Jones, S.N., Hancock, A.R., Vogel, H., Donohower, L.A., Bradley, A., 1998. Overexpression of mdm2 in mice reveals a p53-independent role for mdm2 in tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* 95, 15608–15612.
- Kubbutat, M.H., Jones, S.N., Vousden, K.H., 1997. Regulation of p53 stability by Mdm2. *Nature* 387, 299–303.
- Landers, J.E., Cassel, S.L., George, D.L., 1997. Translational enhancement of mdm2 oncogene expression in human tumor cells containing a stabilized wild type p53 protein. *Cancer Res.* 57, 3562–3568.
- Leach, F.S., Tosino, T., Meltzer, P., Iruel, M., Oliner, J.D., Smith, S., Hill, D.E., Sidransky, D., Kinzler, K.W., Vogelstein, B., 1993. p53 mutation and MDM2 amplification in human soft tissue sarcomas. *Cancer Res.* 53, 2231–2234.
- Levin, B., 1994. The apparatus for nuclear splicing. In: Levin, B. (Ed.), *Gene V*. Oxford Univ. Press, Oxford, pp. 914–940.
- Libri, D., Marie, J., Brody, E., Fuszman, M.Y., 1989. A subfragment of the beta tropomyosin gene is alternatively spliced when transfected into differentiating muscle cells. *Nucleic Acids Res.* 17, 6449–6462.
- Lundgren, K., Morais de Oca Luna, R., McNeill, Y.B., Emswiler, E.P., Spencer, B., Burfield, C.R., Lozano, G., Rosenberg, M.P., Finlay, C.A., 1997. Targeted expression of MDM2 uncouples S phase from mitosis and inhibits mammary gland development independent of p53. *Genes Dev.* 11, 714–725.
- Martin, K., Treusch, D., Hagmann, C., Sorensen, T.S., La Thangue, N.B., Kourilsky, T., 1995. Stimulation of E2F1/DP1 transcriptional activity by MDM2 oncoprotein. *Nature* 375, 691–694.
- Masamoto, R., Tada, M., Nozaki, M., Zhang, C.L., Sawamura, Y., Abe, H., 1998. Short alternative splice transcripts of the mdm2 oncogene correlate to malignancy in human astrocytic neoplasms. *Cancer Res.* 58, 609–613.
- Mosmann, J., Jung, D., Wilczynski, S., Niland, J., 1998. The MDM2 gene amplification database. *Nucleic Acids Res.* 26, 3453–3459.
- Momand, J., Zambetti, G.P., Olson, D.C., George, D., Levine, A.J., 1992. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* 69, 1237–1245.
- Montes de Oca Luna, R., Tabor, A.D., Eberhard, H., Hulbur, D.L., Wirth, L.L., Colman, M.S., Finlay, C.A., Lozano, G., 1996. The organization and expression of the mdm2 gene. *Genomics* 33, 352–357.
- Mueller, A., Oltz, R., Jenkins, T.D., Shaheshtari, A., Nakagawa, H., Inamoto, T., Rustgi, A.K., 1997. A transgenic mouse model with cyclin D1 overexpression results in cell cycle, epidermal growth factor receptor, and p53 abnormalities. *Cancer Res.* 57, 5542–5549.
- Nigro, J.M., Cho, K.R., Fearson, E.R., Kern, S.E., Ruppert, J.M., Oliver, J.D., Kinzler, D.W., Vogelstein, B., 1991. Scrambled exons. *Cell* 64, 607–613.
- Oliner, J.D., Kiezel, K.W., Meltzer, P.S., George, D.L., Vogelstein, B., 1992. Amplification of a gene encoding a p53-associated protein in human sarcomas [see comments]. *Nature* 358, 80–83.
- Oliner, J.D., Pieknap, J.A., Thangalingam, S., Gyuris, J., Kinzler, K.W., Vogelstein, B., 1993. Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53. *Nature* 362, 857–860.
- Puls, R., Anderson, P., 1993. mRNA surveillance by the *Caenorhabditis elegans* smg genes. *Genes Dev.* 7, 1885–1897.
- Sigalas, I., Calvert, A.H., Anderson, J.J., Neal, D.E., Luoco, J., 1996. Alternatively spliced mdm2 transcripts with loss of p53 binding domain sequences: transforming ability and frequent detection in human cancer. *Nat. Med.* 2, 912–917.
- Smith, C.W.J., Porro, E.B., Patton, J.G., Nacal-Ginsard, B., 1989. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature* 342, 243–247.
- Steinman, H., Burstein, E., Lerner, C., Gosselin, J., Pihau, G., Duckett, C., Jones, S., 2004. An alternative splice form of Mdm2 induces p53-independent cell growth and tumorigenesis. *J. Biol. Chem.* 279, 4877–4886.
- Xiao, Z.X., Chen, J., Levine, A.J., Modjtahedi, N., Xing, J., Sellers, W.R., Livingston, D.M., 1995. Interaction between the retinoblastoma protein and the oncoprotein MDM2. *Nature* 375, 694–698.

## Is "Junk" DNA Mostly Intron DNA?

Gane Ka-Shu Wong,<sup>1,3</sup> Douglas A. Passey,<sup>1</sup> Ying-zong Huang,<sup>1</sup> Zhiyong Yang,<sup>1</sup> and Jun Yu<sup>1,2</sup><sup>1</sup>Human Genome Center, Department of Medicine, University of Washington, Seattle, Washington 98195, USA; <sup>2</sup>Human Genome Center, Institute of Genetics, Chinese Academy of Sciences, Beijing, China

Among higher eukaryotes, very little of the genome codes for protein. What is in the rest of the genome, or the "junk" DNA, that, in *Homo sapiens*, is estimated to be almost 97% of the genome? Is it possible that much of this "junk" is intron DNA? This is not a question that can be answered just by looking at the published data, even from the finished genomes. One cannot assume that there are no genes in a sequenced region, just because no genes were annotated. We introduce another approach to this problem, based on an analysis of the cDNA-to-genomic alignments, in all of the complete or nearly-complete genomes from the multicellular organisms. Our conclusion is that, in animals but not in plants, most of the "junk" is intron DNA.

Among higher eukaryotes, very little of the genome codes for protein. What is in the rest of the genome, or the "junk" DNA, that, in *Homo sapiens*, is estimated to be almost 97% of the genome? If a region is gene-poor, is that because there are vast deserts of intergenic DNA between adjacent genes, or is that because the few genes that are there are large, with enormous introns?

First, a few definitions are needed. We consider only the euchromatic portion of the genome. The heterochromatic portion (e.g., centromeres and telomeres) is highly repetitive and largely devoid of genes. It is extremely difficult to clone, extremely polymorphic, and unlikely to be sequenced correctly anytime soon. We define the exons and introns as "intragenic" and everything else as "intergenic." This is not to imply that intergenic DNA is nonfunctional, especially as we have incorporated the promoters into our definition. However, promoters are difficult to identify, whereas exons and introns are reliably identified by cDNA-to-genomic alignments. Lastly, we will use the term "genomic length" to indicate the sum of the exons and introns in a given gene and "cDNA length" to indicate the sum of only the exons.

Even after a genome is completely sequenced, it is not a straightforward matter to determine the intergenic fraction. Indeed, any assessment that is based only on the fraction of the genome that has not been identified by the gene annotations is likely to be an overestimate of the underlying reality. Consider how the genes are annotated. Most current procedures (The *Caenorhabditis elegans* Sequencing Consortium 1998; Dunham et al. 1999; Lin et al. 1999; Mayer et al. 1999; Adams et al. 2000; Hattori et al. 2000) employ a combination of EST/cDNA/protein alignments and ab initio

to exon-prediction programs. Given the incomplete state of the EST/cDNA/protein data, most of the annotated exons are in fact based on the exon-prediction programs, even if parts of certain genes are confirmed by the experimental data. There are two problems (Burset and Guigo 1996; Reese et al. 2000). One is that the exon-prediction programs cannot identify untranslated non-coding exons (i.e., the UTRs). The second, more important, issue is that these programs are not particularly proficient at identifying large genes. There are three reasons: (1) The signal-to-noise ratio can be as low as 1/1000, for the extreme case of a 100-bp exon juxtaposed next to a 100-kb intron; (2) the data sets used to train these programs tend to under-represent the large difficult-to-sequence genes; and (3) the codon-usage statistics, by which the exons are initially identified, are not as informative for the large genes of certain organisms (Wright 1990).

The extent of the large-genes problem is organism dependent. The determinant is the distribution of genomic lengths. If the genomic lengths are distributed over many orders of magnitude, then failure to annotate even a small fraction of the largest genes will leave a much larger fraction of the genome unannotated. In this scenario, there is a critical difference between the following two seemingly similar quantities: the fraction of the genes in the genome that is correctly identified and the fraction of the genome sequence that is labeled as intragenic. The first quantity is far more likely to be correct than the second. It is possible that the total gene counts is essentially correct, while, at the same time, the intragenic fraction is significantly underestimated and the intergenic fraction is significantly overestimated. Indeed, this is precisely the problem for the animal genomes.

Our solution is to determine the distribution of genomic lengths entirely from cDNA-to-genomic alignments (i.e., independent of the exon-prediction

<sup>3</sup>Corresponding author.

E-MAIL: gks@u.washington.edu; FAX (206) 685-7344.

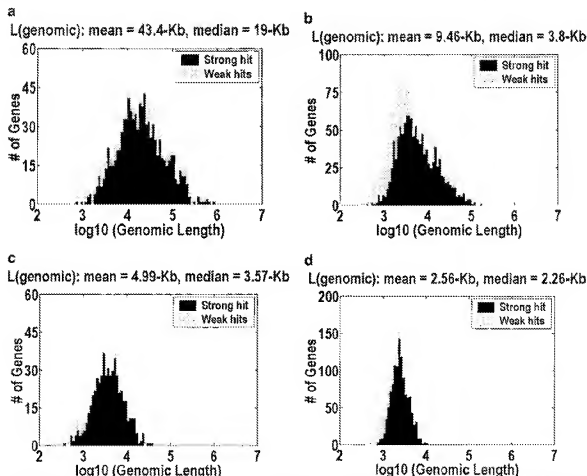
Article and publication are at [www.genome.org/cgi/doi/10.1101/gr.146900](http://www.genome.org/cgi/doi/10.1101/gr.146900)

programs). Then, compare the mean genomic length to the mean gene-to-gene distance. The former is taken from the cDNA alignments, but the latter is computed as the ratio of the euchromatic genome size, divided by the gene count, taken from the annotations. Reliable results are expected for *Drosophila melanogaster* and *Caenorhabditis elegans*, because genome sequencing for these organisms is complete and estimates of the gene-to-gene distance are available. For *Arabidopsis thaliana*, the published chromosomes (Lin et al. 1999; Mayer et al. 1999) agree to 4.5%, so we can safely extrapolate to the entire genome. In contrast, for *H. sapiens*, the published chromosomes (Dunham et al. 1999; Hattori et al. 2000) differ by 243%, reflecting the heterogeneity in the gene densities of warm-blooded vertebrates (Bernardi 2000). Coupled with the difficulties of determining the mean genomic length, a result of the lack of

large genomic contigs, we refer extensively to the model organism results to guide our interpretations of the *H. sapiens* data.

## RESULTS

Figure 1 depicts the distribution of genomic lengths for *H. sapiens*, *D. melanogaster*, *C. elegans*, and *A. thaliana*. Table 1 is a numerical summary. The animal distributions span 2–3 orders of magnitude, but the plant distribution spans only one order of magnitude. The implication for the large genes problem can be estimated by considering how many of the largest genes would have to be unidentified for half of the intragenic space to be missing. The figures range from 11% and 10% at one extreme, in *H. sapiens* and *D. melanogaster*, to 30% at the other extreme, in *A. thaliana*. Furthermore, the only organism in which the intergenic fraction is



**Figure 1** Distribution of genomic lengths for (a) *Homo sapiens*, (b) *Drosophila melanogaster*, (c) *Caenorhabditis elegans*, and (d) *Arabidopsis thaliana*. Dark shading indicates strong hits. Weak hits (lightly shaded) represent cDNA-to-genomic alignments with <3 exons or <30% of the cDNA length aligned. An overwhelming majority of these weak hits are actually complete alignments with only one or two exons. Instances in which <50% of the cDNA is aligned represent 7.3%, 3.3%, 1.2%, and 0.9% of the genes in the four organisms, respectively.

**Table 1. Estimated Intergenic Fractions**

	<i>Homo sapiens</i>	<i>Drosophila melanogaster</i>	<i>Caenorhabditis elegans</i>	<i>Arabidopsis thaliana</i>
Suchromatin	3180000	123000	97800	130600
Sequenced DNA	369000	123000	91600	119000
Gene-to-gene	45.4	9.0	5.3	4.7
cDNA aligned	1061	1628	583	1401
Genomic quality	1.2	23.3	2.4	15.7
Twisted genes	6%	8%	4%	1%
95 Percentile	2.5	0.9	0.8	0.9
Genomic length	43.4	9.5	5.0	2.6
95 Percentile	165.5	36.3	14.2	5.4
% missing half	11%	10%	21%	30%
Intergenic DNA	Discussed in text of article	3%	10%	46%

The first three rows list the suchromatin genome size, the amount of genomic sequence that was analyzed, and the annotation-based estimate of the gene-to-gene distance. The next three rows describe the cDNA alignments. These rows list the number of aligned cDNAs, our quality assessment for the genomic contigs (i.e., the median of the genomic contig size divided by the genomic length for the 95th-percentile gene), and our estimate of the frequency of nested genes (i.e., genes on the reverse strand or inside an intron). The genomic length is given in the next three rows by its arithmetic mean, and its 5th or 95th percentile values. Next, we indicate what fraction of the largest genes would have to be unidentified for half of the intragenic space to be missing. The last row lists the intergenic fraction, computed by correcting the mean genomic length for nested genes, dividing that by the mean gene-to-gene distance, and subtracting the result from one. Note: in *Drosophila melanogaster*, we do not count scaffold pairs longer than 1 kb as contiguous when computing the genomic quality. All lengths are reported in kb.

greater than 10% is *A. thaliana*, even though we have included the minor correction for nested genes (genes on the reverse strand or inside an intron). This correction is computed by counting the occurrences of nested genes in our cDNA alignments, and adjusting for the fact that we do not detect every such occurrence because we do not have all of the cDNAs.

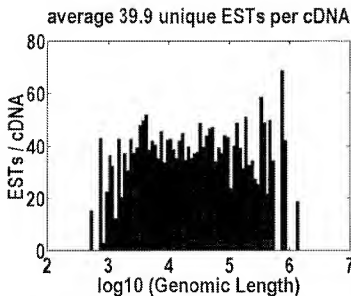
The main uncertainty in our method is that we must extrapolate from a subset of the genes to the entire genome to determine the mean genomic length. There will be sampling biases, but they can be categorized and subcategorized as follows: (1) the extent to which cDNA data are enriched for large or small genes, (2) the extent to which genomic data are biased for large or small genes, and then, are the gene-rich regions done first by sequencing projects? Are the contigs large enough for us to align the large genes?

We will argue that the problem is primarily in the genomic data, not the cDNA data. Furthermore, to the extent that there are sampling biases, the tendencies are always to underestimate the mean genomic length and to overestimate the intergenic fraction.

There are two reasons to suspect that biases in the cDNA data will cause us to underestimate the mean genomic length. Keep in mind that large genes are highly correlated with large cDNAs (this paper; data not shown). The first explanation is that full-length cDNAs are extremely difficult to clone, given the ease with which RNA molecules are degraded and the intrinsic bias in the cloning system for smaller inserts.

The second reason is that large RNA molecules require more time to transcribe, so large genes might be less highly expressed and more difficult to isolate. However, this expectation is incorrect, because the transcription machinery operates in parallel. As a measure of the expression levels, in *H. sapiens*, we aligned the 1,856,102 ESTs in GenBank against our cDNA data. Multiple reads from the same clone were counted only once. Figure 2 shows that there is no significant variation in EST coverage as a function of genomic length. Notice that the normalization procedures (Fuller et al. 1996) applied to the EST libraries do not affect the rare transcripts, in which we were looking for an effect. The conclusion is that cDNA data, extracted from GenBank, can be representative of all genomic lengths.

Genomic data are biased in two ways. First, there is a sociologic bias toward sequencing gene-rich regions first. Second, even when a genome is complete, lack of long-range contiguity, on the scale of the largest genes, will reduce the estimate of the mean genomic length, because any breaks in the alignment are most likely to occur across the largest introns. Both issues are relevant in the *H. sapiens* data. In Figure 3, we demonstrate that the aligned data are biased toward GC-rich genes, which are of smaller genomic lengths (Bernardi 2000). As for contiguity, we estimate the extent of the problem by computing the ratio of the median genomic contig size to the genomic length of the 95th percentile gene. Ideally, this ratio would be much greater than one. Table 1 shows that it is much greater



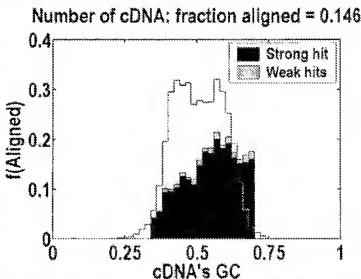
**Figure 2** Is the collection of *Homo sapiens* cDNA sequence biased? We aligned the 1,856,162 ESTs in GenBank to our cDNA sequences and plotted the number of aligned ESTs as a function of the genomic length. Multiple reads from the same clone are counted only once. There is no obvious bias, indicating that cDNAs for genes of every genomic length are equally easy to isolate.

than one in *D. melanogaster* and *A. thaliana*. It is only moderately greater than one in *C. elegans*, but that is less important for this organism, because the genomic lengths are not as broadly distributed. However, in *H. sapiens*, the ratio is 1.2, and it would have been even smaller had we not used genomic data from a new division of GenBank in which all of the overlapping clones have been joined together (Jiang et al. 1999).

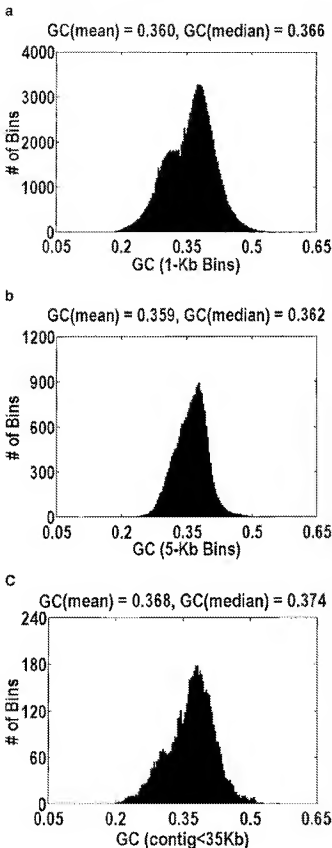
We can estimate the severity of these biases with the different versions of the *D. melanogaster* genomic data. Specifically, we repeated the alignments with the same cDNA data but switched to the 34.9 Mb of finished clone-by-clone genomic data that was available prior to the completion of the whole-genome shotgun (Adams et al. 2000). The contig quality measure is then 2.8, and the resultant mean genomic length of 7.1 kb is off the mark by 34%. By comparing those cDNAs aligned in both data sets, we find that 16% of this effect is attributable to the contig problem. The other 16% is attributable to the bias toward sequencing gene-rich regions first. An even more dramatic example of these biases is *Mus musculus*, which has a contig quality measure of 0.3 and a mean genomic length

of 9.7 kb. If we assume that there is no difference between *M. musculus* and *H. sapiens*, this estimate is off the mark by 447%. Parenthetically, another unreliable way to estimate the mean genomic length is to extract GenBank annotations. The annotated genes in that 34.9 Mb of genomic data for *D. melanogaster* have a mean genomic length of 3.0 kb, which is off the mark by 317%.

The essential conclusion is that our 43.4 kb figure for the mean genomic length in *H. sapiens* is a substantial underestimate, even if it is already 10 times larger than the training set used for these exon-prediction programs. However, the gene count itself is also uncertain. The traditional estimate of 70,000 (Antequera and Bird 1993; Fields et al. 1994) has recently been challenged by substantially lower estimates, from 35,000 to 45,000 (Ewing and Green 2000; Haffner et al. 2000; Roesi-Crollius et al. 2000). How can we interpret the *H. sapiens* data? If we accept the traditional gene count of 70,000, our mean genomic length of 43.4 kb predicts an intergenic fraction of 10%. Suppose we inflate our estimate by the same 34% discrepancy that was observed between the two *D. melanogaster* data sets. The gene count that would be consistent with the same 10% intergenic fraction is



**Figure 3** Is the collection of *Homo sapiens* genomic sequence biased? We computed the probability that cDNAs of a particular GC content aligned to genomic sequence, given that only 369 Mb of nonredundant finished genomic sequence were available. The solid line (on an arbitrary scale) indicates the initial collection of cDNAs. The obvious bias toward GC-rich cDNAs is important because these are known to correspond to smaller genes (Bernardi 2000). Dark shading shows strong hits; light shading shows weak hits.



then 51,400. Considering that the contig quality is much worse in *H. sapiens* than in the clone-by-clone *D. melanogaster* data, it is likely that the mean genomic length is underestimated by >34%. Thus, the gene count would have to be substantially less than the current low estimates of 35,000 to 45,000 for our arguments to allow much intergenic DNA.

Given the uncertainty in our method, we cannot give a precise estimate for the intergenic fraction in *H. sapiens*. However, we are prepared to argue that the intergenic fraction in *H. sapiens* cannot be as large as it is for *A. thaliana*, because, at such a high intergenic fraction, the distribution of GC content for genomic DNA is bimodal, as in Figure 4. Fitting the data to a sum of Gaussians reveals that the main mode is centered at 0.382, which is almost identical to the 0.390 GC content of the aligned *A. thaliana* genes. The relative ratio of the two modes implies an intergenic fraction of 30%, which is smaller than the 46% estimate derived from genomic length arguments but not unexpectedly so, because some of the intergenic DNA could have a GC content that is similar to the intragenic DNA. The reason why this bimodality has not been reported previously is that it is extremely sensitive to how the data are plotted. Specifically, the histogram bins must be smaller than the mean genomic length, and smaller genomic contigs (i.e., those sequenced because they contain a likely gene) cannot be used. That said, no such bimodality is observed in *H. sapiens*, *D. melanogaster*, or *C. elegans*, regardless of how the data are plotted.

## DISCUSSION

So why do most genome annotation efforts continue to report so much intergenic DNA? One of the most conspicuous features of the recent annotations for *H. sapiens* chromosomes 21 and 22 is the small handful of megabase-sized regions

**Figure 4** Distribution of GC content for anonymous genomic sequence in *Arabidopsis thaliana*. The idea that a significant fraction of the genome is intergenic, coupled with the fact that intergenic DNA has a lower GC content than intragenic DNA, suggests that this distribution will be bimodal. However, the bimodality is easily obscured by how the data are plotted. *a* and *b* differ in the size of the bins over which the GC content is computed, 1 kb and 5 kb, respectively. Bin sizes larger than the average gene size of 2.6 kb obscure the effect because every bin is likely to contain a mixture of intragenic and intergenic DNA. *a* and *c* differ in the genomic contigs that are plotted (every contig or only contigs <35 kb, respectively). By removing the large-insert clones favored by the genome centers, what is left behind are those sequences that were analyzed only because they contain a likely gene. Hence, the bimodality disappears.



with absolutely no annotated genes. In all likelihood, each of these regions has one or more large genes, with no counterpart in the EST/cDNA/protein data and which are not being detected by the exon-prediction programs. After accounting for large genes, the remainder of the presently unannotated regions will likely be attributed to untranslated non-coding exons and flanking introns. We must reiterate that the fraction of the genes that is missing does not have to be large to explain away most of the unannotated regions.

What is important is not the precise intergenic fraction or the precise gene count but, at the risk of extrapolating from a limited number of genomes, the differences between plants and animals. There is evidence that plant and animal genomes are organized in different ways. In *H. sapiens*, large genes are caused by a combination of large introns and more introns per gene (this paper; data not shown). At least 35.4% of the total length of the introns in our *H. sapiens* data is due to interspersed repeats (e.g., *Alu* and *L1*). The true fraction is undoubtedly greater, as older repeats, whose sequences are >50% diverged from the ancestral consensus, cannot be identified by existing methods (Srnfl 1996). Analysis of orthologous genes in *Fugu rubripes* and *H. sapiens* reveals that much of the 10-fold difference in the sizes of these two genomes can be explained by differences in intron sizes (Ngar et al. 1996). In contrast, analysis of syntenic loci among grasses reveals that much of the 40-fold difference in the size of these genomes can be explained by their extensive repeat-filled intergenic regions (SanMiguel et al. 1996; Bennetzen et al. 1998).

The conclusion is that, in animals, most repeats integrate into intron DNA, but, in plants, most repeats integrate into intergenic DNA. Is there something different about the nature of the repeats that insert into animals and plants? Does this dichotomy reflect differences in the operation of the introns and promoters? The answers to these questions will be critical for our understanding of the evolution of large-scale genome features.

## METHODS

In *H. sapiens*, cDNA data were extracted from GenBank release 112, but genomic data were downloaded, at the same time, from the ucw division for nonredundant joined-contigs (Jung et al. 1999). In *D. melanogaster*, cDNA data were taken from release 115 (Jung/15/1999), but genomic data were taken from the whole-genome shotgun (Duhms et al. 2000). In *C. elegans* and *A. thaliana*, both cDNA and genomic data were extracted from release 115.

For the cDNA-to-genomic alignments, we required a 98% base pair agreement. We examined the intron sequences for the consensus splice sites, GT-AC, but we also accepted as a substitute GC-AG, albeit in <1% of the data. Weak bases, defined as those with <3 exons or <50% of the cDNA length aligned, were plotted separately to verify that they were not anomalous.

Introns systems-related (cDNAs [i.e., with the description immunoglobulin, Ig, HLA, MHC, V-region, etc.] were removed. Other redundancies were eliminated, up front by removing all cDNAs that are 90% contained in some other cDNA and post-alignment by comparing the genomic coordinates of the aligned exons. Raw genomic lengths were extrapolated to compensate for incomplete alignments—a small correction even for *H. sapiens*, where a total of 86% of the cDNA lengths was aligned. As another quality control, we required that the exact coordinates of the coding region (i.e., the open reading frame) be known, even though it reduced the number of genes in our final data set.

The partial alignment correction is done by computing an adjusted number of exons,  $N_{\text{exons}}^{\text{adj}}$ , with a linear extrapolation. The adjusted genomic length,  $L_{\text{genomic}}^{\text{adj}}$ , is  $N_{\text{exons}}^{\text{adj}} \times L_{\text{exon}}^{\text{avg}} + (N_{\text{exons}} - N_{\text{exons}}^{\text{adj}}) \times L_{\text{intron}}^{\text{avg}}$ , is extrapolated in a similarly linear manner, with the averages  $L_{\text{exon}}^{\text{avg}}$  and  $L_{\text{intron}}^{\text{avg}}$  being defined on a per gene basis. Because non-coding terminal exons are generally larger than coding interior exons, both extrapolations are only performed across the coding portion of the cDNA sequence. The intention is to ensure that, if anything, we underestimate the mean genomic length.

## ACKNOWLEDGMENTS

We thank Phil Green, Maynard Olson, Carl Toni, and Lee Rosen for many useful discussions and suggestions. This work was partly supported by a grant from the National Institutes of Health (1 R01 ES09909).

The publication costs of this article were delayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

## REFERENCES

- Adams, M.D., Schöler, S.H., Holt, R.A., Evans, C.A., Gwinn, J.E., et al. 2000. The genome sequence of *Arabidopsis thaliana*. *Science* **287**: 2185–2195.
- Antequera, F., and Sudd, A.P. 1993. Number of CpG islands and genes in human and mouse. *Proc. Natl. Acad. Sci. USA* **90**: 11995–11999.
- Bennetzen, J.L., SanMiguel, P., Chen, M., Khurana, A., Francis, M., et al. 1998. Grass genomes. *Proc. Natl. Acad. Sci. USA* **95**: 1978–1979.
- Bernardi, G. 2000. Isochores and the evolutionary genomics of vertebrates. *Gene* **241**: 3–17.
- Buser, M., and Gupta, R. 1985. Evaluation of gene structure prediction programs. *Genomics* **3**: 553–567.
- The C. elegans Sequencing Consortium. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**: 2012–2018.
- Duqueno, L., Mizuno, N., Rice, S.A., Chubb, S., Hunt, A.R., et al. 1999. The DNA sequence of human chromosome 22. *Nature* **402**: 489–495.
- Ngar, G., Sandwell, R., Aparicio, S., Maerz, A., Yenilal, B., et al. 1996. Small is beautiful: Comparative genomics with the pufferfish (*Fugu rubripes*). *Trans. Genet.* **12**: 148–150.
- Swing, B., and Goep, P. 2003. Analysis of annotated sequence tags indicates 35,640 human genes. *Nat. Genet.* **25**: 232–234.
- Peck, C., Adams, M.D., Venter, O., and Foster, J.C. 1994. *Bradyrhizobium* genes in the human genome? *Nat. Genet.* **7**: 345–348.
- Hattori, M., Duggan, A., Taylor, T.D., Watanabe, H., Yada, Y., et al. 2000. The DNA sequence of human chromosome 21. *Nature* **405**: 311–319.
- Miller, L.D., Larnson, G., Becker, M., Banbura, M.F., Chappell, B., et al. 1996. Generation and analysis of 280,000 human expressed sequence tags. *Genome Res.* **6**: 807–814.
- Jong, W., Chen, H., Skothe, H., and Schink, G.D. 1999. Making

- effective use of human genomic sequence data. *Trends Genet.* **15**: 264–266.
- Lin, N., Kaul, S., Rounsley, S., Shiu, J.P., Benito, M.L., et al. 1999. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402**: 761–768.
- Mayer, K., Schaller, C., Wambutt, R., Murphy, G., Volkwein, G., et al. 1999. Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. *Nature* **402**: 769–777.
- Reese, M.C., Hartzell, G., Harris, N.L., Ophir, U., Abul, J.P., and Lewis, S.E. 2000. Genome annotation assessment in *Drosophila melanogaster*. *Genome Res.* **10**: 483–501.
- Roeck, C., Hillier, B., Jaffe, O., Searles, A., Davila, C., Bonavent, J., et al. 2000. Estimate of human gene number provided by genome-wide analysis using *Tetrahymena nigroviridis* DNA sequence. *Nat. Genet.* **25**: 235–236.
- Saint-Macq, F., Tikhonov, A., Jin, Y.E., Mochizuki, N., Zakharen, O., et al. 1996. Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**: 765–766.
- Smith, A.E. 1996. The origin of interspersed repeats in the human genome. *Hum. Opin. Genet. Dev.* **6**: 733–748.
- Wright, I. 1991. The 'effective number of codons' used in a gene. *Gene* **87**: 23–29.

Received May 23, 2006; accepted in revised form August 29, 2006.

## RELATED PROCEEDINGS APPENDIX

None